PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT

INTERNATIONAL ATTECATION PUBLIS	UED	JAPER THE PATENT COOPERATIO	JN IREALI (PCI)
(51) International Patent Classification 5:		(11) International Publication Number:	WO 94/22477
A61K 39/12, 39/00, C07K 17/00, 3/00, 13/00, 15/00, C07H 15/12, 13/00	A1	(43) International Publication Date:	13 October 1994 (13.10.94)
(21) International Application Number: PCT/US (22) International Filing Date: 25 March 1994 (European patent (AT, BE, CH,	L. CN, JP, RU, UA, US, DE, DK, ES, FR, GB, GR,
(30) Priority Data: 08/037,816 26 March 1993 (26.03.93)	ţ	Published With international search report	
(60) Parent Application or Grant (63) Related by Continuation US 08/037.8	01 <i>4 (C</i> 1	m. l	
US 08/037,8 Filed on 26 March 1993 (
(71) Applicant (for all designated States except US): PRO PHARMACEUTICALS, INC. [US/US]; Old Saw M Road, Tarrytown, NY 10591 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): HASEL, R [CH/US]; 1713 Baldwin Road, Yorktown Heig 10598 (US). MADDON, Paul, J. [US/US]; A #25C, 60 Haven Avenue, New York, NY 10032 (ghts, N Apartme	Y	
(74) Agent: WHITE, John, P.; Cooper & Dunham, 30 Ro Plaza, New York, NY 10112 (US).	•		
(54) Title: HIV-1 VACCINES, ANTIBODY COMPOSIT USES THEREOF	IONS I	ELATED THERETO, AND THERAPEUTI	C AND PROPHYLACTIC
(57) Abstract			
The invention provides a recombinant nucleic acid m comprising the mutant HIV-1 envelope glycoprotein, antibo	olecuk odies at	which encodes a mutant HIV-1 gp120 envel d methods of treating individuals.	ope glycoprotein, vaccines
·			
			i

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MIR	5.0 1. 1
ΑÜ	Australia	GE	Georgia		Mauritania
BB	Barbados	· GN	Guinea	MW	Malawi
BE	Belgium	GR	Greece	NE	Niger
BF	Burkina Faso	HU		NL	Netherlands
BG	Bulgaria		Hungary	NO	Norway
BJ	Benin	Œ	Ireland	NZ	New Zealand
BR	Brazil	n	Italy	PL	Poland
BY	_	JP	Japan	PT	Portugal
	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic	SD	Suden
CG	Congo		of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	
CM	Cameroon	ш	Liechtenstein	SN	Slovakia
CN	China	LK	Sri Lanka		Senegal
cs	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV		TG	Togo
DE	Сетрасу		Latvie	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
ES		MD	Republic of Moldova	UA	Ukraine
FI	Spain	MG	Madagascar	US	United States of America
	Finland	MI.	Mali	UZ	Uzbekistan
FR	France	MIN	Mongolia	VN	Vict Nam
GA	Gabon				

HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES THEREOF

Background of the Invention

10

Throughout this application, various publications are referenced by Arabic numerals. Full citations for these references may be found at the end of the specification immediately preceding the claims. The disclosure of these publications is hereby incorporated by reference into this application to describe more fully the art to which this invention pertains.

The life cycle of animal viruses is characterized by a 20 series of events that are required for the productive infection of the host cell. The initial step in the replicative cycle is the attachment of the virus to the cell surface, which attachment is mediated by the specific interaction of the viral attachment protein (VAP) to 25 receptors on the surface of the target cell. The differential pattern of expression of these receptors is largely responsible for the host range and tropic properties of viruses. In addition, an effective immune response against many viruses is mediated through neutralizing 30 antibodies directed against the VAP. The interaction of the VAP with cellular receptors and the immune system therefore plays a critical role in infection and pathogenesis of viral disease.

The human immunodeficiency virus type 1 (HIV-1) infects primarily helper T lymphocytes, dendritic cells, and monocytes/macrophages--cells that express surface CD4--leading to a gradual loss of immune function. This loss of function results in the development of the human acquired

immunodeficiency syndrome (AIDS) (1). The initial phase of the HIV-1 replicative cycle involves the high-affinity interaction between the HIV-1 exterior envelope glycoprotein gp120 and cell surface CD4 (K_d approximately 4 x 10.9 M) (2). 5 Several lines of evidence demonstrate the requirement of this interaction for viral infectivity. The introduction into CD4 human cells of cDNA encoding CD4 is sufficient to render otherwise resistant cells susceptible to HIV-1 infection (3). In vivo, viral infection appears to be 10 restricted to cells expressing CD4, indicating that the cellular tropism of HIV-1 is largely determined by the pattern of cellular expression of CD4. Following the binding of HIV-1 gp120 to cell surface CD4, viral and target membranes fuse by a mechanism that is poorly 15 understood, resulting in the introduction of the viral capsid into the target cell cytoplasm (4).

Mature CD4 has a relative molecular mass (Mr) of 55 kDa and consists of an N-terminal 372-amino acid extracellular domain containing four tandem immunoglobulin-like regions (V1-V4), followed by a 23-amino acid transmembrane domain and a 38-amino acid cytoplasmic segment (5, 6). In experiments using truncated sCD4 proteins, it has been shown that the determinants for high-affinity binding to HIV-1 gp120 lie solely within the N-terminal immunoglobulin-like domain (V1) (7-9). Mutational analysis of V1 has defined a discrete binding site (residues 38-52) that comprises a region structurally homologous to the second complementarity-determining region (CDR2) of immunoglobulin genes (9).

The production of large quantities of sCD4 has permitted a structural analysis of the two N-terminal immunoglobulin-like domains (V1V2). The structure determined at 2.3

angstrom resolution reveals that the molecule has two tightly-associated domains, each of which contains the immunoglobulin-fold connected by a continuous beta strand. The putative binding sites for monoclonal antibodies, class II major histocompatibility complex (MHC) molecules, and HIV-1 gp120, as determined by mutational analyses, map on the molecular surface (10, 11).

The HIV-1 envelope gene env encodes an envelope glycoprotein precursor, gp160, which is cleaved by cellular proteases before transport to the plasma membrane to yield gp120 and gp41. The membrane-spanning glycoprotein, gp41, is non-covalently associated with gp120, a purely extracellular glycoprotein. The mature gp120 molecule is heavily glycosylated (approximately 24 N-linked oligosaccharides), contains approximately 480 amino acid residues with 9 intrachain disulfide bonds (12), and projects from the viral membrane as a dimeric or multimeric molecule (13).

Mutational studies of HIV-1 gp120 have delineated important functional regions of the molecule. The regions of gp120 that interact with gp41 map primarily to the N- and C-termini (14). The predominant strain-specific neutralizing epitope on gp120 is located in the 32-34 amino acid residue third variable loop, herein referred to as the V3 loop, which resides near the center of the gp120 sequence (15). The CD4 binding site maps to discontinuous regions of gp120 that include highly conserved or invariant amino acid residues in the second, third, and fourth conserved domains (the C2, C3, and C4 domains) of gp120 (16). It has been postulated that a small pocket formed by these conserved residues within gp120 could accommodate the CDR2 loop of CD4, a region defined by mutational analyses as important in interacting with gp120 (17).

HIV-1 gp120 not only mediates viral attachment to surface CD4 molecules, but also serves as the major target of antibodies which neutralize non-cell-associated virus and inhibit cell to cell viral transmission.

5

There are two major classifications of HIV-1-neutralizing antibodies: type-specific and group-common (15). specific neutralizing antibodies primarily recognize linear determinants in the highly variable V3 loop of gp120. These 10 antibodies act by inhibiting fusion between HIV-1 and the target cell membrane, and generally neutralize only a particular isolate of, or closely related strains of, HIV-1. Sequence variation within the V3 loop, as well as outside of this region, permits viruses to escape neutralization by 15 anti-V3 loop antibodies. In contrast, group-common neutralizing antibodies primarily recognize discontinuous or conformational epitopes in gp120, and possess the ability to neutralize a diverse range of HIV-1 isolates. These broadly neutralizing antibodies often recognize a site on gp120 20 which overlaps the highly conserved CD4-binding site, and thus inhibits gp120-CD4 binding.

A structural relationship has been demonstrated between the V3 loop and the C4 region of gp120 which region constitutes both part of the CD4 binding site and part of the conserved neutralization epitopes. It was observed that deleting the V3 loop resulted in significantly increased binding of a panel of broadly neutralizing hMoAbs (neutralizing human monoclonal antibodies) to the CD4 binding site (18).

30

A major goal in AIDS vaccine development is to develop a vaccine able to protect a subject against the numerous genetic variants of HIV-1 that infect humans. Although cell-mediated immune responses might serve to control infection in HIV-1-infected individuals, several lines of

5

evidence demonstrate that protection against infection is mainly mediated by neutralizing antibodies directed against Early experiments showed that immunization of chimpanzees with recombinant gp120 induced a protective 5 immune response against challenge with the homologous HIV-1 strain (17). This protection correlated with the presence of high-titer neutralizing antibodies against the V3 loop of In addition, passive immunization of chimpanzees with a V3-loop neutralizing monoclonal antibody resulted in 10 protection against challenge with the homologous HIV-1 strain (19). Although protection against challenge was demonstrated in these two experiments, recent studies have questioned the clinical relevance of these findings. example, these neutralizing antibodies recognize the V3 loop 15 determinants of a single strain, and not conserved or discontinuous epitopes. Thus, these antibodies lack the ability to neutralize the broad spectrum of HIV-1 strains present in an HIV-1 population. Furthermore, the challenge virus was the homologous HIV-1 laboratory adapted LAI (HTLV-20 IIIB) strain and not one of the primary isolates that contain considerable gp120 sequence heterogeneity. Since these experiments showed that gp120 subunit vaccination induces an immune response effective against only the homogeneous HIV-1 strain used as an antigen, it is unlikely 25 that the vaccination regimens used in these studies would be useful in humans.

Individuals infected by HIV-1 typically develop antibodies that neutralize the virus in vitro, and neutralization titers decrease with disease progression (19). Analysis of sera from HIV-1-infected humans indicates that type-specific neutralizing antibodies appear early in infection. Later in the course of infection, a more broadly neutralizing antibody response develops. However this antibody response is of significantly lower titer and/or affinity.

Fractionation studies of HIV-1 antibody-positive human sera reveal that the type-specific neutralizing activity is primarily directed against linear determinants in the V3 loop of gp120 (20). There was no correlation found among antibodies between the ability to neutralize divergent HIV-1 isolates and reactivity to the V3 loop of these isolates. In contrast, the broadly neutralizing antibodies present in HIV-1 antibody-positive human sera primarly recognize discontinuous epitopes in gp120 which overlap the CD4-binding site and block gp120-CD4 binding. In other words, the broadly neutralizing activity of neutralizing antibodies is not merely the result of additive anti-V3 loop reactivities against diverse HIV-1 isolates which appear during infection.

6

15

Recently, several groups have generated human monoclonal antibodies (hMoAbs) derived from HIV-1 infected individuals which possess type-specific or group-common neutralizing activities (17). The type-specific neutralizing hMoAbs were found to recognize linear determinants in the V3 loop of gp120. In contrast, the group-common neutralizing hMoAbs generally recognize discontinuous epitopes which overlap the CD4-binding site and block gp120-CD4 binding.

- The V3 loop is a highly immunodominant region of gp120 which partially interacts with the CD4-binding region. The presence of the V3 loop region on gp120 may skew the humoral immune response away from producing antibodies which specifically bind to the CD4-binding domain of gp120.
- Furthermore, the advantages of removing the V3 loop to expose the CD4-binding domain of gp120 to the immune system would be countered by the fact that the exposed CD4-binding site would still have a high affinity for cell surface CD4. In other words, a mutant gp120 protein missing only the V3
- 35 loop would quickly bind to CD4+ cells and would thus be

7

hampered in generating an immune response against the exposed CD4-binding site.

The subject invention provides a mutant HIV-1 gp120 envelope glycoprotein which overcomes both the problems of V3 loop immunodominance and of the high affinity to CD4. The subject invention further provides vaccines comprising the mutant HIV-1 gp120 envelope glycoprotein, antibodies which specifically bind to the CD4-binding site of HIV-1 gp120 envelope glycoprotein, pharmaceutical compositions comprising these antibodies, and methods of using these vaccines and compositions to treat or prevent HIV-1 infection.

Summary of the Invention

The subject invention provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain_(W->X) point mutation, wherein X is an amino acid residue other than tryptophan. In the preferred embodiment, X is a valine residue.

- 10 In one embodiment, the nucleic acid molecule is a DNA molecule. The DNA molecule may be a plasmid. In one embodiment, the plasmid comprises the sequence of the plasmid designated PPI4-tPA.
- 15 In one embodiment, the C4 domain is an HIV-1_{LAI} gp120 envelope glycoprotein C4 domain. The mutant HIV-1 gp120 envelope glycoprotein may be a mutant HIV-1_{LAI} gp120 envelope glycoprotein.
- In another embodiment, the C4 domain is an HIV-1 $_{\rm R-FL}$ gp120 envelope glycoprotein C4 domain. The mutant HIV-1 gp120 envelope glycoprotein may be a mutant HIV-1 $_{\rm R-FL}$ gp120 envelope glycoprotein.
- 25 The subject invention also provides the mutant HIV-1 gp120 envelope glycoprotein encoded by the recombinant nucleic acid molecule of the subject invention.
- The subject invention further provides a vaccine which comprises a therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.

The subject invention further provides a method of treating

an HIV-1-infected subject, which comprises immunizing the HIV-1-infected subject with the vaccine of the subject invention, thereby treating the HIV-1-infected subject.

9

- 5 The subject invention further provides a vaccine which comprises a prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.
- The subject invention further provides a method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the HIV-1-exposed subject's becoming infected with HIV-1.

The subject invention further provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the non-HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the non-HIV-1-exposed subject's becoming infected with HIV-1.

The subject invention further provides a method of obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein, which method comprises (a) immunizing a non-HIV-1-exposed subject with the vaccine of the subject invention, (b) recovering from the immunized subject serum comprising said antibodies, and (c) partially purifying said antibodies, thereby obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein. In the preferred embodiment, the subject is a human.

The subject invention further provides the partially purified antibodies produced by the method of the subject invention.

10

- 5 The subject invention further provides a pharmaceutical composition, which comprises a therapeutically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.
- The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.

The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1 in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.

The subject invention further provides a composition which comprises a prophylactically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.

The subject invention further provides a method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises administering to the HIV-1-exposed subject a dose of the composition of the subject invention effective to reduce the population of HIV-1 in the HIV-1-exposed subject, thereby reducing the likelihood of the subject's becoming infected with HIV-1.

In one embodiment, the subject is a medical practitioner. In another embodiment, the subject is a newborn infant.

Finally, the subject invention provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1 as a result of exposure thereto during an incident wherein there is an increased risk of exposure to HIV-1, which comprises administering to the subject immediately prior to the incident a dose of the composition of the subject invention effective to reduce the population of HIV-1 to which the subject is exposed during the incident, thereby reducing the likelihood of the subject's becoming infected with HIV-1. In one embodiment, the subject is a medical practitioner.

Brief Description of the Figures

Figure 1

depicting the boundaries of the five constant domains (C1-C5) and the five variable domains (V1-V5). The amino acid residue numbering above the box begins at the initiator methionine found at the beginning of the signal sequence (S) and is approximated based on a consensus of all known HIV-1 gp120 amino acid sequences. Also shown are the C4 domain amino acid sequences of HIV-1 strains LAI and JR-FL. Above the C4 domain sequences are indicated two mutations that reduce gp120 binding to cell surface CD4; tryptophan to valine and aspartate to alanine.

15

Figure 2

PPI4-tPA-gp120LAI. Expression vector with the HIV-1LAI gp120
gene fused to the CMV MIE promoter, and the tPA signal
sequence replacing the HIV-1 gp120 signal sequence.
20 Abbreviations: CMV MIE = cytomegalovirus major immediate
early, E = enhancer, P = promoter, EXA = Exon A, INA =
Intron A, EXB = Exon B, tPA ss = human tissue plasminogen
activator signal sequence, gp120 = glycoprotein 120, BGH =
bovine growth hormone, AMP = ampicillin resistance gene, and
DHFR = dihydrofolate reductase gene.

Figure 3

CMV MIE promoter fused to tPA-gp120_{LAI}. The nucleotide sequence of the CMV MIE promoter/enhancer region is shown fused to the HIV-1_{LAI} gp120 gene that contains the tPA signal sequence. The numbering of nucleotide sequence begins with the HincII site and the numbering of the amino acid sequence begins with the first methionine found in the tPA signal sequence. The tPA signal sequence is fused in-frame to Thr₃₁

of gp120, the first amino acid found in mature gp120. The signal sequence is shown in bold as are various landmark restriction sites used for cloning as discussed in the text. The locations of Exon A, Intron A, Exon B and the transcription start site and the signal cleavage site are indicated.

Figure 4

Transient expression of gp120. Autoradiograph of ³⁵S-labeled supernatants from COS cell transfectants, immunoprecipitated with a CD4-immunoglobulin-Protein A-Sepharose complex, and run on a reducing 10% SDS-PAGE gel. The plasmids used for transfection were: Lane 1: Mock transfected cells; lane 2: a vector encoding a CD4-immunoglobulin chimera as a positive transfection control; lane 3: PPI4-tPA-gp120_{IAI}; and lane 4: PPI4-tPA-gp120_{IRFL}. Positions of molecular weight markers are indicated.

Figure 5

Determination of gp120 concentration by ELISA. Panel A: Concentrations of gp120 in media of CHO cell lines, stably transfected with PPI4-tPA-gp120_{LAI}, determined by ELISA. Panel B: A standard curve was established using known amounts of gp120.

25

Figure 6

Expression of gp120 in stably transfected CHO cells. Autoradiograph of ³⁵S-labeled supernatants from stable CHO cell lines, immunoprecipitated with a CD4-immunoglobulin30 Protein A-Sepharose complex, and run on a reducing 10% SDS-PAGE gel. Lane 1: clone 9; lane 2: clone 13; lane 3: clone 6; lane 4: Clone 5. Positions of molecular weight markers are indicated.

Figure 7

tPA-gp120_{IR-FL}. The nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1_{IR-FL} gp120 is shown. The NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg₃₅ and Val₃₆ is indicated.

Figure 8

tPA-gp120_{LAI}-V3⁽⁾. The nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1_{LAI} gp120 with the V3 loop deleted and replaced with the pentapeptide TGAGH is shown. The V3 loop replacement and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg₃₅ and Thr₃₆ is indicated.

Figure 9

tPA-gp120_{JR-FL}-V3^(*). The nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1_{JR-FL} gp120 with the V3 loop deleted and replaced with the pentapeptide TGAGH is shown. The V3 loop replacement and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg₃₅ and Val₃₆ is indicated.

Figure 10

25

tPA-gp120_{LAI}-V3⁽³⁾-CD4⁽³⁾. Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1_{LAI} gp120, with the V3 loop deleted and replaced with the pentapeptide TGAGH, and Trp403 mutated to Val. The mutations and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg33 and Thr36 is indicated.

Figure 11

tPA-gp120_{JR-FL}-V3^(*)-CD4^(*). Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1_{JR-FL} gp120, with the V3 loop deleted and replaced with the pentapeptide TGAGH, and Trp₃₉₆ mutated to Val. The mutations and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg₃₅ and Val₃₆ is indicated.

10

Figure 12

tPA-gpl20_{LAI}-CD4⁽⁾. Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1_{LAI} gpl20. The Trp₄₃₇ to Val CD4 binding mutation, the NarI and NotI restriction endonuclease sites used for cloning, and the predicted site of cleavage by signal peptidase between Arg₃₅ and Thr₃₆ are shown in bold.

Figure 13

20 <u>tPA-gp120_{JR-FL}-CD4^(*)</u>. Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1_{JR-FL} gp120. The Trp₄₂₄ to Val CD4 binding mutation, the NarI and NotI restriction endonuclease sites used for cloning and the predicted cleavage by signal peptidase between Arg₃₅ and Val₃₆
25 are shown in bold.

Figure 14

Expression of gp120 in stably transfected CHO cells.

Autoradiograph of super ³⁵S-labeled supernatants from stable CHO cell lines, immunoprecipitated with MoAb F105-Protein A-Sepharose complex, and run on a reducing 10% SDS-PAGE gel. Panel A: Lane 1: tPA-gp120_{LAI} CHO cells; lane 2: tPA-gp120_{LAI}-V3⁽⁺⁾ CHO cells; lane 3: tPA-gp120_{LAI}-V3⁽⁺⁾ CD4⁽⁺⁾ CHO cells. Panel B: Lane 1: tPA-gp120_{IR-FL} CHO cells; lane 2: tPA-gp120_{IR-FL}-V3⁽⁺⁾

CHO cells; lane 3: $tPA-gp120_{JR.FL}-V3^{(\cdot)}-CD4^{(\cdot)}$ CHO cells. Positions of molecular weight markers are indicated.

Figure 15

5 Purified gp120 proteins.

Silver stained 10% SDS-PAGE gel with a sample of purified gp120 proteins. Panel A: Lane 1: tPA-gp120_{LAI} CHO cells; lane 2: tPA-gp120_{LAI}-V3^(·) CHO cells; lane 3: tPA-gp120_{LAI}-V3^(·)-CD4^(·) CHO cells. Panel B: Lane 1: tPA-gp120_{IR-FL} CHO cells; lane 2: tPA-gp120_{IR-FL}-V3^(·) CHO cells; lane 3: tPA-gp120_{IR-FL}-V3^(·)-CD4^(·) CHO cells. Positions of molecular weight markers are indicated.

Figure 16

Analysis of binding of recombinant mutant gp120 to cell surface human CD4 by FACS.

Plate 1. DG44 cells, a subclone of CHO cells which lack expression of the human CD4 protein, were used as control. Increasing concentrations of HIV-1 gp120_{LAI} did not show an increase in specific fluoresence when compared to background. Plate 2. DG44 #3 cells are a CHO cell line transfected with the cDNA clone encoding the human CD4 protein. Increasing concentrations of HIV-1 gp120_{LAI} show a dramatic increase (or shift) in fluoresence. Plate 3. Similar to Plate 2 but the HIV-1 gp120_{LAI}-V3^(*) protein was added. Again a large shift indicating binding to the DG44 #3 cells was seen. Plate 4. DG44 #3 cells were incubated with either HIV-1 gp120_{LAI}-V3^(*)-CD4^(*) protein or MoAb OKT4A an antibody with high affinity for human CD4. Only OKT4A bound to the cells.

Detailed Description of the Invention

The plasmids designated PPI4-tPA-gp120_{LAI} and PPI4-tPA-gp120_{JR}.

FL were deposited pursuant to, and in satisfaction of, the
requirements of the Budapest Treaty on the International
Recognition of the Deposit of Microorganisms for the
Purposes of Patent Procedure with the American Type Culture
Collection (ATCC), 12301 Parklawn Drive, Rockville, Maryland
20852 under ATCC Accession Nos. 75431 and 75432,
respectively. The plasmids PPI4-tPA-gp120_{LAI} and PPI4-tPAgp120_{JR-FL} were deposited with the ATCC on March 12, 1993.

The subject invention provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain_(W->X) point mutation, wherein X is an amino acid residue other than tryptophan. In the preferred embodiment, X is a valine residue.

- 20 In one embodiment, the nucleic acid molecule is a DNA molecule. The DNA molecule may be a plasmid. In one embodiment, the plasmid comprises the sequence of the plasmid designated PPI4-tPA.
- The V3 loop of HIV-1 gp120 envelope glycoprotein is shown in Figure 1. The V3 loop is demarcated by cysteine residues at both its N- and C-termini. As used herein, a V3 loop deletion means a deletion of one or more amino acid residues between the terminal cysteine residues, with the proviso that there must be three or more amino acid residues situated between the two terminal cysteine residues in a V3 loop deletion. These three or more amino acid residues may either be residues originally present in the V3 loop, or exogenous residues. For example, as shown in the

Experimental Details section <u>infra</u>, the pentapeptide TGAGH is situated between the two terminal cysteine residues. Variations in the size of the V3 loop deletion illustrated herein are tolerable without affecting the overall structure of the mutant HIV-1 gp120 envelope glycoprotein, as is well known to those skilled in the art.

As used herein, "C4 domain" means the HIV-1 gp120 envelope glycoprotein C4 domain having the following consensus 10 sequence:

 $\begin{array}{l} X_{1}X_{2}X_{3}CX_{4}IX_{5}X_{6}X_{7}X_{8}X_{9}X_{10}WX_{11}X_{12}X_{13}X_{14}X_{15}AX_{16}YX_{17}X_{18}-\\ PX_{19}X_{20}X_{21}X_{22}X_{23}X_{24}X_{25}X_{26}SX_{77}X_{28}TGX_{29}X_{30}X_{31}X_{32}RX_{33}GX_{34}, \end{array}$

15 wherein X₁ = T, I, V, K or R; X₂ = L, I or H; X₃ = P, Q, L or
T; X₄ = R, K or G; X₅ = K or E; X₆ = Q or E; X₇ = F, I or V;
X₈ = I, V or M; X₉ = N, R or K; X₁₀ = M, R, L or T; X₁₁ = Q, R
or V; X₁₂ = E, K, G, R, V or A; X₁₃ = V, T, A or G; X₁₄ = G or
E; X₁₅ = K, R, E, or Q; X₁₆ = M, V, I or L; X₁₇ = A, T or D; X₁₈
20 = P or L; X₁₉ = I or F; X₂₀ = S, R, G, K, N, A, E or Q; X₂₁ =
G or R; X₂₂ = Q, L, P, N, K, V, T, E or I; X₂₃ = I, V or L; X₂₄
= R, K, S, N, G, I, T, E or I; X₂₅ = C or R; X₂₆ = S, L, I, T,
P, E, V, K, D or N; X₂₇ = N, K or L; X₂₈ = I or V; X₂₉ = L, P
or I; X₃₀ = L or I; X₃₁ = L or I; X₃₂ = T, A, I, V or E; X₃₃ =
25 D or E; X₃₄ = G or V.

The C4 domain consensus sequence is based on existing C4 domain sequence information from various HIV-1 strains, and thus is not necessarily an exhaustive consensus sequence.

The conserved tryptophan residue shown in bold after residue X_{10} is the only conserved tryptophan residue in the C4 domain. As used herein, a C4 domain $_{(W->X)}$ point mutation is a mutation of the above-identified conserved C4 domain tryptophan residue to an amino acid residue other than

19

tryptophan. For example, a C4 domain $(w_{-}>v)$ point mutation is a mutation of the conserved C4 domain tryptophan residue to a valine residue.

In one embodiment, the C4 domain is an HIV-1_{LAI} gp120 envelope glycoprotein C4 domain. The sequence of the HIV-1_{LAI} gp120 C4 domain is: TLPCRIKQFINMWQEVGKAMYAPPISGQIRCS-SNITGLLLTRDGG. The mutant HIV-1 gp120 envelope glycoprotein may be a mutant HIV-1_{LAI} gp120 envelope glycoprotein.

10

In another embodiment, the C4 domain is an HIV-1_{JR-FL} gp120 envelope glycoprotein C4 domain. The sequence of the HIV-1_{JR} FL gp120 C4 domain is: TLPCRIKQIINMWQEVGKAMYAPPIRGQIRCS-SNITGLLLTRDGG. The mutant HIV-1 gp120 envelope glycoprotein may be a mutant HIV-1_{JR-FL} gp120 envelope glycoprotein.

HIV-1_{LM} is a laboratory-adapted strain that is tropic for phytohemagglutinin (PHA) -stimulated peripheral lymphocytes (PBLs) and immortalized human T-cell lines. In 20 contrast, $HIV-1_{IR-FL}$ was isolated from brain tissue taken at autopsy that was co-cultured with lectin-activated normal human PBLs. $\text{HIV-1}_{JR\text{-FL}}$ is tropic for PHA-stimulated PBLs and blood-derived macrophages but will not replicate in transformed T-cell lines. Mutant HIV-1 gp120 envelope 25 glycoproteins derived from a clinical isolate of HIV-1 such as JR-FL may possess new or different epitopes compared to the laboratory-adapted HIV-1 strains that are beneficial for successful vaccination. Although only the HIV-1 $_{
m LAI}$ and HIV- $\mathbf{1}_{\mathbf{R}.\mathbf{FL}}$ strains are used herein to generate the mutant HIV-1 30 gp120 envelope glycoproteins of the subject invention, other HIV-1 strain could be substituted in their place as is well known to those skilled in the art.

The V1 and V2 variable regions of gp120 are unnecessary for

CD4 binding (21). Therefore the mutant HIV-1 gp120 envelope glycoprotein of this invention can either include or exclude the V1 and V2 variable regions.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain_(Asp->X) point mutation, wherein the aspartate residue is between amino acid residues X₁₅ and X₁₆ in the C4 consensus sequence, and X is an amino acid residue other than aspartate or glutamate. In the preferred embodiment, X is an alanine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain_(Gh->X) point mutation, wherein the glutamate residue is between amino acid residues X_{15} and X_{16} in the C4 consensus sequence, and X is an amino acid residue other than aspartate or glutamate. In the preferred embodiment, X is an alanine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1_{LAI} gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 domain_(ep378->X) point mutation, wherein X is an amino acid residue other than aspartate or glutamate. In the preferred embodiment, X is a lysine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1_{R-FL} gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 domain_(sep369->X) point mutation, wherein X is an amino acid residue other than aspartate or glutamate. In the preferred

embodiment, X is a lysine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1LAI gp120 5 envelope glycoprotein comprising a V3 loop deletion and a C3 $domain_{(glu380->X)}$ point mutation, wherein X is an amino acid residue other than glutamate. In the preferred embodiment, X is a glutamine residue.

10 The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant $HIV-1_{JR-FL}$ gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 $\operatorname{domain}_{(glu371->X)}$ point mutation, wherein X is an amino acid residue other than glutamate. In the preferred embodiment,

15 X is a glutamine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant $HIV-1_{LAI}$ gp120 envelope glycoprotein comprising a V3 loop deletion and a C2 20 domain $_{(th/267->X)}$ point mutation, wherein X is an amino acid residue other than threonine. In the preferred embodiment, X is an arginine residue.

The subject invention additionally provides a recombinant 25 nucleic acid molecule which encodes a mutant $HIV-1_{JR.FL}$ gp120 envelope glycoprotein comprising a V3 loop deletion and a C2 $domain_{(b)250->x)}$ point mutation, wherein X is an amino acid residue other than threonine. In the preferred embodiment, X is an arginine residue.

30

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising (a) a V3 loop deletion, or (b) a one of the C2, C3 or C4 domain point mutations discussed supra.

The point mutations in the recombinant nucleic acid molecules described <u>supra</u> are selected based on their ability to reduce the affinity of the mutant gp120 glycoprotein encoded thereby for CD4. As used herein, the term "reduce the affinity" means to reduce the affinity by at least two-fold.

One skilled in the art would know how to make recombinant nucleic acid molecules which encode mutant HIV-1 gp120 envelope glycoproteins comprising a V3 loop deletion and the specific C2, C3 or C4 domain point mutations corresponding to those mutations exemplified in the HIV-1_{IR-FL} and HIV-1_{LAI} strains, supra. Furthermore, one skilled in the art would know how to use these recombinant nucleic acid molecules to obtain the proteins encoded thereby, and practice the therapeutic and prophylactic methods of using same, as described herein for the recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain_(W->X) point mutation.

The subject invention also provides the mutant HIV-1 gp120 envelope glycoprotein encoded by the recombinant nucleic acid molecule of the subject invention.

In accordance with the invention, numerous vector systems for expression of the mutant HIV-1 gp120 envelope glycoprotein may be employed. For example, one class of vectors utilizes DNA elements which are derived from animal viruses such as bovine papilloma virus, polyoma virus, adenovirus, vaccinia virus, baculovirus, retroviruses (RSV, MMTV or MoMLV), Semliki Forest virus or SV40 virus.

23

Additionally, cells which have stably integrated the DNA into their chromosomes may be selected by introducing one or more markers which allow for the selection of transfected host cells. The marker may provide, for example, prototropy 5 to an auxotrophic host, biocide resistance, antibiotics) or resistance to heavy metals such as copper or the like. The selectable marker gene can be either directly linked to the DNA sequences to be expressed, or introduced into the same cell by cotransformation. Additional elements 10 may also be needed for optimal synthesis of mRNA. elements may include splice signals, as well transcriptional promoters, enhancers, and termination The cDNA expression vectors incorporating such elements include those described by Okayama (22).

15

The vectors used in the subject invention are designed to express high levels of mutant HIV-1 gp120 envelope glycoproteins in cultured eukaryotic cells as well as efficiently secrete these proteins into the culture medium. The targeting of the mutant HIV-1 gp120 envelope glycoproteins into the culture medium is accomplished by fusing in-frame to the mature N-terminus of the mutant HIV-1 gp120 envelope glycoprotein the tissue plasminogen activator (tPA) prepro-signal sequence.

25

The mutant HIV-1 gp120 envelope glycoprotein may be produced by a) transfecting a mammalian cell with an expression vector for producing mutant HIV-1 gp120 envelope glycoprotein; b) culturing the resulting transfected mammalian cell under conditions such that mutant HIV-1 gp120 envelope glycoprotein is produced; and c) recovering the mutant HIV-1 gp120 envelope glycoprotein so produced.

Once the expression vector or DNA sequence containing the constructs has been prepared for expression, the expression

24

vectors may be transfected or introduced into an appropriate mammalian cell host. Various techniques may be employed to achieve this, such as, for example, protoplast fusion, calcium phosphate precipitation, electroporation or other conventional techniques. In the case of protoplast fusion, the cells are grown in media and screened for the appropriate activity. Expression of the gene encoding a mutant HIV-1 gp120 envelope glycoprotein results in production of the mutant glycoprotein.

10

Methods and conditions for culturing the resulting transfected cells and for recovering the mutant HIV-1 gp120 envelope glycoprotein so produced are well known to those skilled in the art, and may be varied or optimized depending upon the specific expression vector and mammalian host cell employed.

In accordance with the claimed invention, the preferred host cells for expressing the mutant HTV-1 gp120 envelope glycoprotein of this invention are mammalian cell lines. Mammalian cell lines include, for example, monkey kidney CV1 line transformed by SV40 (COS-7); human embryonic kidney line 293; baby hamster kidney cells (BHK); Chinese hamster ovary-cells-DHFR (CHO); Chinese hamster ovary-cells DHFR (DXB11); monkey kidney cells (CV1); African green monkey kidney cells (VERO-76); human cervical carcinoma cells (HELA); canine kidney cells (MDCK); human lung cells (W138); human liver cells (Hep G2); mouse mammary tumor (MMT 060562); mouse cell line (C127); and myeloma cell lines.

30

Other eukaryotic expression systems utilizing non-mammalian vector/cell line combinations can be used to produce the mutant HIV-1 gp120 envelope glycoproteins. These include, but are not limited to, baculovirus vector/insect cell

expression systems and yeast shuttle vector/yeast cell expression systems.

Methods and conditions for purifying mutant HIV-1 gp120 envelope glycoproteins from the culture media are provided in the invention, but it should be recognized that these procedures can be varied or optimized as is well known to those skilled in the art.

- 10 The subject invention further provides a vaccine which comprises a therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.
- 15 A therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein may be determined according to methods well known to those skilled in the art.
- As used herein, adjuvants include, but are not limited to, alum, Freund's incomplete adjuvant (FIA), Saponin, Quil A, Monophosphoryl lipid A (MPL), and nonionic block copolymers (SAF) such as L-121 (Pluronic; Syntex SAF). In the preferred embodiment, the adjuvant is alum, especially in the form of a thixotropic, viscous, and homogeneous aluminum hydroxide gel. The vaccine of the subject invention may be administered as an oil in water emulsion. Methods of combining adjuvants with antigens are well known to those skilled in the art.
- The subject invention further provides a method of treating an HIV-1-infected subject, which comprises immunizing the HIV-1-infected subject with the vaccine of the subject invention, thereby treating the HIV-1-infected subject.
- 35 As used herein, treating an HIV-1-infected subject with the

vaccine of the subject invention means reducing in the subject either the population of HIV-1 or HIV-1-infected cells, or ameliorating the progression of an HIV-1-related disorder in the subject.

5

As used herein, an "HIV-infected subject" means an individual having at least one of his own cells invaded by HIV-1.

- As used herein, "immunizing" means administering a primary dose of the vaccine to a subject, followed after a suitable period of time by one or more subsequent administrations of the vaccine, so as to generate in the subject an immune response against the CD4-binding region of the mutant HIV-1
- 15 gp120 envelope glycoprotein in the vaccine. A suitable period of time between administrations of the vaccine may readily be determined by one skilled in the art, and is usually in the order of several weeks to months.
- In the preferred embodiment, the dose of vaccine administered is an amount sufficient to deliver to the subject between 10ug and 1mg of the mutant HIV-1 gp120 envelope glycoprotein.
- The subject invention further provides a vaccine which comprises a prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.
- 30 A prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein may be determined according to methods well known to those skilled in the art.

The subject invention further provides a method of reducing the likelihood of an HIV-1-exposed subject's becoming

27

infected with HIV-1, which comprises immunizing the HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the HIV-1-exposed subject's becoming infected with HIV-1.

5

As used herein, the subject's becoming infected with HIV-1 means the invasion of the subject's own cells by HIV-1.

As used herein, reducing the likelihood of a subject's becoming infected with HIV-1 means reducing the likelihood of the subject's becoming infected with HIV-1 by at least two-fold. For example, if a subject has a 1% chance of becoming infected with HIV-1, a two-fold reduction in the likelihood of the subject's becoming infected with HIV-1 would result in the subject's having a 0.5% chance of becoming infected with HIV-1. In the preferred embodiment of this invention, reducing the likelihood of the subject's becoming infected with HIV-1 means reducing the likelihood of the subject's becoming infected with HIV-1 means reducing the likelihood of the subject's becoming infected with HIV-1 by at least ten-fold.

As used herein, an HIV-1-exposed subject is a subject who has HIV-1 present in his body, but has not yet become HIV-1-infected.

25

35

The subject invention further provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the non-HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the non-HIV-1-exposed subject's becoming infected with HIV-1.

As used herein, a non-HIV-1-exposed subject is a subject who does not have HIV-1 present in his body.

28

The subject invention further provides a method of obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein, which method comprises (a) immunizing a non-HIV-1-exposed subject with the vaccine of the subject invention, (b) recovering from the immunized subject serum comprising said antibodies, and (c) partially purifying said antibodies, thereby obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein. In the preferred embodiment, the subject is a human.

As used herein, partially purified antibodies means a composition which comprises antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein, and consists of fewer protein impurities than does the serum from which the anti-CD4-binding domain antibodies are derived. A protein impurity means a protein other than the anti-CD4-binding domain antibodies. For example, the partially purified antibodies might be an IgG preparation.

Methods of recovering serum from a subject are well known to those skilled in the art. Methods of partially purifying antibodies are also well known to those skilled in the art, and include, by way of example, filtration, ion exchange chromatography, and precipitation.

In one embodiment, the partially purified antibodies comprise an immune globulin (IG) preparation. IG can be purified from serum by a two-step process. Initially, serum is fractionated by the cold ethanol method of Cohn, et al. (29). Cohn Fraction II has as its main protein component IgG immunoglobulin present as monomers, dimers and aggregates. Fraction II is then purified to produce IVIG

29

(immune globulin intravenous) using a variety of purification methods which include, for example, ion exchange, DEAE chromatography, acid pH 4.25 diafiltration, PEG precipitation or Pepsin treatment. The final product is stabilized (e.g., glucose + NaCl) and the final IgG concentration is fixed at between about 3% and about 6%.

The subject invention further provides the partially purified antibodies produced by the method of the subject of invention.

The subject invention further provides a pharmaceutical composition, which comprises a therapeutically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.

A therapeutically effective amount of the partially purified antibodies of the subject invention may be determined according to methods well known to those skilled in the art.

20

Pharmaceutically acceptable carriers are well known to those skilled in the art and include, but are not limited to, 0.01-0.1M and preferably 0.05M phosphate buffer or 0.8% Additionally, such pharmaceutically acceptable saline. 25 carriers may be aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include 30 water. alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid 35 and nutrient replenishers, electrolyte replenishers such as

30

those based on Ringer's dextrose, and the like. Preservatives and other additives may also be present, such as, for example, antimicrobials, antioxidants, chelating agents, inert gases and the like.

5

The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.

As used herein, administering may be effected or performed using any of the various methods known to those skilled in the art. The administering may comprise administering intravenously. The administering may also comprise administering intramuscularly. The administering may further comprise administering subcutaneously.

The dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10 mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100 mg/kg and 2g/kg of protein if administered intravenously.

30

The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-

35 1 in the HIV-1-infected subject, thereby treating the HIV-1-

infected subject.

The dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1 in the HIV-1 infected subject may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10 mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100 mg/kg and 2g/kg of protein if administered intravenously.

The subject invention further provides a composition which comprises a prophylactically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.

A prophylactically effective amount of the partially purified antibodies of the subject invention may be determined according to methods well known to those skilled in the art.

The subject invention further provides a method of reducing
the likelihood of an HIV-1-exposed subject's becoming
infected with HIV-1, which comprises administering to the
HIV-1-exposed subject a dose of the composition of the
subject invention effective to reduce the population of HIV1 in the HIV-1-exposed subject, thereby reducing the
likelihood of the subject's becoming infected with HIV-1.

In one embodiment, the subject is a medical practitioner. The medical practitioner may be a medical practitioner exposed to an HIV-1-containing bodily fluid. As used herein, the term "medical practitioner" includes, but is in no way

limited to, doctors, dentists, surgeons, nurses, medical laboratory assistants, and students in health care programs.

32

In another embodiment, the subject is a newborn infant. The newborn infant may be a newborn infant born to an HIV-1-infected mother.

The dose of the composition of the subject invention effective to reduce the population of HIV-1 in the HIV-110 exposed subject may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100 mg/kg and 2g/kg of protein if administered intravenously.

The vaccines and pharmaceutical compositions of the subject invention may also ameliorate the progression of an HIV-1-related disorder in a subject to whom the vaccines or pharmaceutical compositions were administered while the subject was either non-HIV-1-exposed or HIV-1-exposed, but not yet HIV-1-infected.

25

Finally, the subject invention provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1 as a result of exposure thereto during an incident wherein there is an increased risk of exposure to HIV-1, which comprises administering to the subject immediately prior to the incident a dose of the composition of the subject invention effective to reduce the population of HIV-1 to which the subject is exposed during the incident, thereby reducing the likelihood of the subject's becoming infected with HIV-1. In one embodiment, the

33

subject is a medical practitioner.

An incident wherein there is an increased risk of exposure to HIV-1 includes, for example, receiving a blood 5 transfusion, sexual contact with an HIV-1-infected individual, and performing a HIV-1-containing bodily fluid-exposing medical procedure.

As used herein, "immediately prior to the incident" means 10 within one month of the incident. In the preferred embodiment, "immediately prior to the incident" means within one day of the incident.

The dose of the composition of the subject invention effective to reduce the population of HIV-1 to which the subject is exposed during the incident may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100mg/kg and 2g/kg of protein if administered intravenously.

One embodiment of this invention is a method of substantially reducing the likelihood of a non-infected medical practitioner's becoming infected with HIV-1 during a bodily fluid-exposing medical procedure involving a patient, which comprises administering to the patient during a suitable time period an amount of the composition of the subject invention effective to substantially reduce the likelihood of the non-infected medical practitioner's becoming infected with HIV-1 by virtue of contact with the patient's bodily fluid during the medical procedure.

35

34

As used herein, a bodily fluid is any fluid which is present in the human body and is capable of containing infectious HIV-1 in an HIV-1-infected patient. Bodily fluids include, but are not limited to, saliva, cerebrospinal fluid, tears, vaginal secretions, urine, alveolar fluid, synovial fluid and pleural fluid.

Another embodiment of this invention is a method of substantially reducing the likelihood of a non-HIV-1infected newborn infant's becoming infected with HIV-1 prior to or during birth from an HIV-1-infected mother, which comprises administering to the mother prior to birth an amount of the composition of the subject invention effective to substantially reduce the likelihood of the non-HIV-1infected newborn infant's becoming infected with HIV-1 by virtue of contact with the patient's bodily fluid.

In order to facilitate an understanding of the Experimental Details section which follows, certain frequently occurring 20 methods and/or terms are best described in Maniatis et al. (23).

This invention will be better understood by reference to the Experimental Details which follow, but those skilled in the art will readily appreciate that the specific experiments detailed are only illustrative of the invention as described more fully in the claims which follow thereafter.

Experimental Details

Nomenclature

As used herein, V3^(*) indicates a V3 loop deletion from HIV-1 gp120 envelope glycoprotein. As used herein, CD4^(*) indicates a point mutation in the C4 domain of HIV-1 gp120 envelope glycoprotein which mutation inhibits CD4 binding to the mutant HIV-1 gp120 envelope glycoprotein. The structure of HIV-1 gp120 envelope glycoprotein is shown in Figure 1.

10 Materials and Methods

Construction of PPI4-tPA-gp120 expression vector. An expression vector was constructed that consisted of the cytomegalovirus major immediate early (CMV 15 promoter/enhancer linked to the HIV-1 gene, which gene had its signal sequence replaced by the tPA signal sequence. The CMV MIE promoter/enhancer sequences were derived from pSVCC1 (24) consisting of 1580 base pairs of contiguous DNA that is immediately 5' to the initiator ATG. In sequential 20 order, the functional domains of the CMV promoter are: the promoter/enhancer region; a transcriptional initiator site; exon A (a non-coding exon); intron A; and 17 nucleotides of exon B (non-coding sequences). The viral promoter sequences were ligated to a gene construct consisting of the 25 nucleotide sequences encoding amino acids -35 to -1 of human tPA (25) fused in-frame to HIV-1_{LAI} env amino acids 31 through 515, ending with a TGA stop codon. The construction was performed in two parts. The majority of the CMV promoter could be isolated as a 1560 bp Hinc II/Pst I fragment which 30 was ligated to a Pst I/Not I 1590 bp DNA fragment that contained the remainder of the CMV promoter, the initiator ATG, the tPA signal sequence and the mature HIV-14 env protein coding sequence.

The latter fragment was assembled using the polymerase chain reaction as follows. Primer 1 (GATCCTGCAGTCACCGTCCTTGACA-CGATGGATGCAATGAAGAGA) and primer 2 (AAGTCTTCTCCTCGGTCTTGT-CTTTTTAACACCCAG) were used to amplify the nucleic acid 5 sequences encoding the tPA signal sequence amino acids -35 to -1 from plasmid pMAM neo-s (Clonetech), thus producing a 150 bp fragment. A second 1440 bp DNA fragment was amplified using primer 3 (TTCAGAAGAGGAGCCAGAACAGAAAATTGTGGGTC), primer 4 (GGAAAAAAGCGGCCGCTCATTTTCTCTCTCTGCACCACTC), and pENV (26) as a template. The PCR fragments were pooled, desalted, and excess primer removed by ultrafiltration through a centricon-100 unit (Amicon). An aliquot of the pooled material was then subjected to a second round of amplification in the presence of primers 1 and 4 to produce 15 a 1590 bp fragment, which was then digested with Pst I and The CMV promoter fragment and the HIV-1 env Not I. fragment were then ligated together, and the entire transcription unit subcloned into PPI4, which is a eukaryotic shuttle vector that contains an ampicillin 20 resistance gene, an SV40 origin of replication and a DHFR gene whose transcription is driven by the ß-globin promoter. The final construct, PPI4-tPA-gp120LAI, is shown in Figure 2.

The expression vector is then used as the prototype vector

for the expression of gp120 proteins that are derived from other HIV-1 strains or mutated as described in the methods section. The vector was constructed so that unique Nar I and Not I sites flank the gp120 sequence, thus facilitating the removal of the gp120 gene cassette and the subsequent insertion of other gene cassettes (Figure 2).

- Expression of HIV-1_{LAI} gp120 in mammalian cells.
- a. <u>Transient expression</u>.

CosM5 cells grown in DMEM containing 10% fetal calf serum

were split to 75% confluence. On the following day, the cells were transfected for 16-20 hours with 10 micrograms of CsCl-purified PPI4-tPA-gp120_{LAI} DNA by the standard CaPO₄ (5) precipitation technique. After transfection, fresh medium was added to the cells. Analysis of the products synthesized 96-120 hours post-transfection was performed by radiolabelling the transfectants with ³⁵S-cysteine for 12-18 hours, followed by precipitation of media using a CD4-immunoglobulin-Protein A-Sepharose complex, followed by SDS-10 PAGE under reducing conditions (Figure 4).

b. Stable expression.

Dhfr Chinese hamster ovary cells (CHO) were transfected with 20 micrograms of CsCl-purified DNA. Approximately 3-5 15 days post-transfection, cells were placed in selective medium (nucleoside-free alpha MEM containing 10% dialyzed fetal calf serum). Approximately 10-15 days post-selection, individual cell clones were picked. Media was analyzed for gp120 expression by radiolabelling the cells with 35-20 cysteine for 12-18 hours, followed by precipitation of media using a CD4-immunoglobulin-Protein A-Sepharose complex, followed in turn by SDS-PAGE under reducing conditions The levels of gp120 in the media of these (Figure 6). clones were also quantitated (Figure 5) by ELISA performed The method involves coating 96-well plates 25 as follows. overnight with sheep polyclonal IgG against the highly conserved C-terminus of gp120 (D7234, Aalto Bioreagents). After washing, dilutions of a standard gp120 preparation in cell growth medium, or supernatant from the stably-30 transfected cells, were incubated for 1 hour. The plates were washed again, and incubated for one hour with a horseradish peroxidase-conjugated anti-gp120 monoclonal antibody (9204, DuPont). Following a final wash, the peroxidase substrate OPD (DuPont) was added and the amount

WO 94/22477 PCT/US94/03282

38

of gp120 determined by comparing absorbance of unknowns with a standard curve. Standards were prepared from purified gp120 made in CHO cells, a small quantity of which was obtained from Celltech Ltd. Clones expressing the highest levels were subjected to successive rounds of amplification of the newly introduced DNA sequences in increasing concentrations of methotrexate. Stable CHO cell lines were thus generated which secrete at least 1 microgram/milliliter of HIV-1_{IAI} gp120.

10

3. Construction of PPI4-tPA-gp120 P.FL

a. The HIV-1_{LAI} gp120 env nucleotide sequence in PPI4-tPA-gp120_{LAI} was replaced by the nucleotide sequence encoding the mature gp120_{IR-FL} protein. Using the polymerase chain reaction, the JR-FL sequences were amplified from pUC112-1 (27) using primer 5 (GATCGGCGCCAGAGTAGAAAAGTTGTGGGTCAC) and primer 4. The PCR fragment was digested with the restriction endonucleases Nar I and Not I, and the fragment subcloned in between the Nar I and Not I sites in PPI4-tPA-20 gp120_{LAI} to generate PPI4-tPA-gp120_{IR-FL} (Figure 7).

b. Transient expression.

CosM5 cells grown in DMEM containing 10% fetal calf serum were split to 75% confluence. On the following day, the cells were transfected for 16-20 hours with 10 micrograms of CsCl-purified PPI4-tPA-gp120_{IR-PL} DNA by the standard CaPO₄ (5) precipitation technique. After transfection, fresh medium was added to the cells. Analysis of the products synthesized 96-120 hours post-transfection was performed by radiolabelling the transfectants with ³⁵S-cysteine for 12-18 hours, followed by precipitation of media using a CD4-immunoglobulin-Protein A-Sepharose complex, followed by SDS-PAGE under reducing conditions (Figure 4).

4. Construction of PPI4-tPA-gp120, 1-V3(-).

The V3 loop in tPA-gp120 $_{LAI}$ consists of amino acids Cys $_{306}$ through Cys333. In the $V3^{(\cdot)}$ mutant, the amino acids in between these cysteines are replaced by the pentapeptide 5 sequence Thr-Gly-Ala-Gly-His. Using the Transformer Site-Directed Mutagenesis Kit (Clonetech), the V3 loop sequence in PPI4-tPA-gp120LAI is altered using the mutagenic primer 6 (CTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAACATTAGTAGAGC) primer 7 (CTCGAGCATGCATTCGAAGCTCGCTGATC) as a selection 10 primer. Primer 7 changes a unique Xba I site in the backbone of the parent PPI4 plasmid into a unique BstB I Briefly, the mutagenesis method requires incubating of the parent plasmid with the mutagenic primer and the selection primer, denaturing at 100°C for 3 minutes and then 15 chilling on ice. In the presence of buffered deoxynucleotide triphosphates and T4 DNA polymerase, the primers are allowed to initiate the polymerization of one strand of plasmid DNA. T4 DNA ligase is used to seal the newly synthesized DNA strand to form a covalently closed circle. 20 Hybrid plasmids are then transformed into a MutS strain of E. coli that is deficient in mismatch repair. allowing for the growth of transformed cells, DNA is purified from the cells and digested with the selection restriction endonuclease, in this case Xba I. 25 plasmids are cleaved by Xba I while the mutant plasmid remains resistant to cleavage by virtue of the Xba I to BstB I conversion. Digested DNA is then used to transform E. coli, and colonies harboring the mutant plasmid are picked. Multiple mutagenic primers can be used in a single round of The amino acid sequence of the modified 30 mutagenesis. protein is shown in Figure 8.

5. Construction of PPI4-tPA-gp120_{R-FL}-V3(·).

The V3 loop in tPA-gp120 $_{\rm IR-FL}$ consists of amino acids Cys $_{293}$

through Cys₃₇₇. In the V3^(*) mutant, the amino acids in between these cysteines are replaced by the pentapeptide sequence Thr-Gly-Ala-Gly-His. Using the Transformer Site-Directed Mutagenesis Kit (Clonetech), the V3 loop sequence in PPI4-tPA-gp120_{JR-FL} is altered using the mutagenic primer 6 (CTGTAGAAATTGTACAGGTGCTGGACATTGTAACATTAGTAGAGC) and primer 7 as a selection primer. The amino acid sequence of the modified protein is shown in Figure 9.

10 6. Construction of PPI4-tPA-gp120_{IAI}-CD4⁽¹⁾.

Using the Transformer Site-Directed Mutagenesis Kit (Clonetech), the selection primer 7, and the mutagenic primer 8 (CAATTTATAAACATGGTGCAGGAAGTAGG), Trp437 of tPA-gp120_{LAI}, which is in an equivalent position to the tryptophan residue in the HXBc2 strain of HIV-1, is mutated to a Val in the expression vector PPI4-tPA-gp120_{LAI} to generate PPI4-tPA-gp120_{LAI}-CD4⁽⁻⁾. The sequence for gp120_{LAI}-CD4⁽⁻⁾ is shown in Figure 12.

20 7. Construction of PPI4-tPA-gp120_{RFI}-CD4⁽⁴⁾.

In a fashion similar to that described above, Trp_{424} of $tPA-gp120_{IR-FL}$ is mutated to a Val in the expression vector PPI4- $tPA-gp120_{IR-FL}$ using the selection primer 7 and the mutagenic primer 9 (CAAATTATAAACATGGTGCAGGAAGTAGG) to generate PPI4-

tPA-gp120 $_{IR-FL}$ -CD4 $^{(\cdot)}$. The sequence for gp120 $_{IR-FL}$ -CD4 $^{(\cdot)}$ is shown in Figure 13.

8. Construction of PPI4-tPA-qp120_{LAI}-V3(+)-CD4(+).

The tPA-gp120_{LAI} double mutant, V3⁽⁺⁾-CD4⁽⁺⁾, is constructed by including the mutagenic primers 6 and 8, and the selection primer 7 simultaneously in the reaction tube with PPI4-tPA-gp120_{LAI} as the DNA template. The final construct is named PPI4-tPA-gp120_{LAI}-V3⁽⁺⁾-CD4⁽⁺⁾, and its sequence is shown in figure 10.

PCT/US94/03282

9. Construction of PPI4-tPA-gp120_{JR-FL}-V3⁽⁺⁾-CD4⁽⁺⁾.

The tPA-gp120_{JR-FL} double mutant, V3^(*)-CD4^(*), is constructed by including the mutagenic primers 6 and 9, and the selection primer 7 simultaneously in the reaction tube with PPI4-tPA-gp120_{JR-FL} as the DNA template. The final construct is named PPI4-tPA-gp120_{JR-FL}-V3^(*)-CD4^(*), and its sequence is shown in figure 11.

10. Expression of mutant HIV-1 gp120 in mammalian cells.

10 a. <u>Transient expression</u>.

CosM5 cells grown in DMEM containing 10% fetal calf serum are split to 75% confluence. On the next day, the cells are transfected for 16-20 hours with 10 micrograms of CsCl-purified mutant HIV-1 DNA by the standard CaPO₄ (5)

- precipitation technique. After transfection, fresh medium is added to the cells. Analysis of the products synthesized 96-120 hours post-transfection is performed by radiolabelling the transfectants with 35-cysteine for 12-18 hours, followed by precipitation of media using a sheep polyclonal IgG against the highly conserved C-terminus of gp120.
 - b. Stable expression.

Dhfr Chinese hamster ovary cells (CHO) are transfected with 20 micrograms of CsCl-purified DNA encoding the native or mutant HIV-1 gp120 glycoproteins. Approximately 3-5 days post-transfection, cells are placed in selective medium (nucleoside-free alpha MEM containing 10% dialyzed fetal calf serum). Approximately 10-15 days post-selection, 30 individual cell clones are picked. Media is analyzed for

gp120 expression by radiolabelling the cells with ³⁵S-cysteine for 12-18 hours, followed by quantitative immunoprecipitation of media using a sheep polyclonal IgG against the highly conserved C-terminus of gp120, followed

turn by SDS - PAGE under reducing conditions. Alternatively, one can quantitate the level of gp120 by ELISA performed as follows. The method involves coating 96well plates overnight with sheep polyclonal IgG against the 5 highly conserved C-terminus of gp120 (D7234, Bioreagents). After washing, dilutions of a standard gpl20 preparation in cell growth medium, or supernatant from the stably-transfected cells, are incubated for 1 hour. plates are washed again, and incubated for one hour with a 10 human MoAb (F105, AIDS Research & Reference Reagent Program, No. 857). The plates are washed again, and incubated again for 1 hour with a horseradish-peroxidase-conjugated goat Following a final wash, the anti-human IgG (Cappel). peroxidase substrate OPD (DuPont) is added and the amount of 15 gp120 determined by comparing absorbance of unknowns with a standard curve. Standards are prepared from purified gp120 made in CHO cells, a small quantity of which is obtained from Celltech Ltd. Clones expressing the highest levels are subjected to successive rounds of amplification of the newly 20 introduced DNA sequences in increasing concentrations of methotrexate. Stable CHO cell lines are thus generated which secrete at least 1 microgram/milliliter of mutant HIV-1 gp120.

25 11. Purification of HIV-1 gp120 proteins.

A one-step immunoaffinity procedure is used to purify the recombinant gp120 molecules described. Briefly, culture supernatant is collected and clarified by centrifugation. An immunoaffinity column consisting of a matrix coupled to a sheep polyclonal anti-gp120 IgG (D7234, Aalto Bioreagents) directed against the highly conserved C-terminal end (APTKAKRRVVQREKR) of gp120 is used to specifically adsorb gp120 from the cell culture media. This antisera recognizes native gp120, the V3 loop deletion mutants, and the CD4(4)

mutants since the C-terminal ends of these molecules remain unaltered. The bound gp120 is then eluted with 2M MgCl₂, concentrated by Amicon filtration, and dialyzed into 10 mM HEPES, pH 7.0. The purity of the proteins is determined by SDS-PAGE and silver staining.

12. Characterization of recombinant HIV-1 gp120 proteins. The purified glycoproteins are subjected to extensive biochemical and immunologic characterization. The integrity of the proteins is monitored by SDS-PAGE and silver staining under reducing and non-reducing conditions. The glycoproteins are deglycosylated by treatment with the enzyme N-glycosidase F which cleaves N-linked oligosaccharides, and are assayed by SDS-PAGE and silver staining to monitor molecular weight shifts. The purified glycoproteins are also tested for reactivity with several well characterized anti-gp120 monoclonal antibodies that recognize both linear and discontinuous epitopes. The binding affinity to sCD4 is estimated using an ELISA assay.

The purified proteins HIV-1 gp120_{IAI}, gp120_{IAI}-V3⁽⁺⁾, gp120_{IAI}-V3⁽⁺⁾
)-CD4⁽⁺⁾, gp120_{IR-FL}, gp120_{IR-FL}-V3⁽⁺⁾, and gp120_{IR-FL}-V3⁽⁺⁾, were tested for their ability to bind cell surface human CD4. DG44 #3 cells, a recombinant cell line designed to express human CD4 on the membrane surface, were grown in T flasks and trypsinized. 5 X 10⁵ cells/experiment were aliquoted into FACS buffer (PBS + 2% BSA and 0.1% NaN₃), washed several times in the same buffer, and then incubated with 100 ul of a solution of purified gp120 protein at 5ug/ml in FACS buffer at 37°C for 2 hr. The cells were washed in FACS buffer, and then incubated in 100 ul solution containing 5ug/ml sheep polyclonal IgG against the highly conserved C-terminus of gp120 in FACS buffer at 37°C for 2 hr. The

cells were washed in FACS buffer then incubated in 100 ul

WO 94/22477 PCT/US94/03282

44 .

solution containing FITC-labeled rabbit anti-sheep IgG polyclonal antibody at 37°C for 2 hr. The cells were washed with FACS buffer and then resuspended in 500 ul FACS buffer. The cells were then analyzed on a Becton Dickinson FACScan according to the manufacturer's instructions. As a control for expression of CD4 on the DG44 #3 cells, FITC-labeled OKT4A (Becton Dickinson) was used.

13. A protocol for inoculation of animals with the mutant HIV-1 gpl20 envelope glycoproteins.

Alum is used as an adjuvant during the inoculation series. The inoculum is prepared by dissolving the mutant HIV-1 gp120 envelope glycoprotein antigen in physiologic saline at a final antigen concentration of 100 ug/ml. Preformed alum (aluminum hydroxide gel) is added to the solution to a final level of 500 ug/ml aluminum. The antigen is allowed to adsorb onto the alum gel for two hours at room temperature. Following adsorption, the gel with the antigen is washed twice with physiologic saline and resuspended in the saline to a protein concentration of 100 ug/ml.

Monkeys and/or Guinea Pigs are individually inoculated with four 100 ug doses of the mutant HIV-1 gp120 envelope glycoprotein antigen adsorbed onto alum. Each dose is injected intramuscularly. The doses are delivered one or five months apart (week 0, 4, 8 and 28). the animals are bled at intervals of two or four weeks. Serum samples are prepared from each bleed to assay for the development of specific antibodies as described in the subsequent sections.

30

10

14. Analysis of sera for anti-mutant HIV-1 gp120 envelope glycoprotein IgG antibodies.

Each serum sample is analyzed by ELISA. Polystyrene microtiter plates are coated with 0.5 ug per well of pure 35 mutant HIV-1 gp120 envelope glycoprotein in phosphate-

buffered physiological saline (PBS) at 4°C. Each well is then washed with PBS containing 0.5% TWEEN-20 (PBS-TW). Test serum, diluted serially in PBS-TW, is added to the mutant HIV-1 gp120 envelope glycoprotein-containing wells and allowed to react with the adsorbed mutant HIV-1 gp120 envelope glycoprotein for one hour at 37°C. The wells are then washed extensively in PBS-TW. Each well then receives 0.1% p-nitrophenyl phosphate in 10% diethanolamine, pH 9.8, containing 0.5 mM MgCl₂.6H₂0. The ensuing reaction is allowed to proceed at room temperature for 30 minutes, at which time it is terminated by the addition of 3.0 N NaOH.

The greater the interaction of antibodies in the test serum with the mutant HIV-1 gp120 envelope glycoprotein, the greater is the amount of alkaline phosphatase bound onto the well. The phosphatase enzyme mediates the breakdown of pnitrophenyl phosphate into a molecular substance which absorbs light at a wavelength of 405 nm. Hence, there exists a direct relationship between the absorbance at 405 nm of light at the end of the ELISA reaction and the amount of mutant HIV-1 gp120 envelope glycoprotein-bound antibody. All animals inoculated with mutant HIV-1 gp120 envelope glycoprotein whose serum reacts specifically with the mutant HIV-1 gp120 envelope glycoprotein in the ELISA have a positive antibody response against mutant HIV-1 gp120 envelope glycoprotein.

15. Analysis of sera for activity which specifically neutralizes HIV-1 infectivity.

Virus-neutralizing activity is determined with an assay based on the use of multiplicity curves in which the ratio of infectious virus surviving antibody treatment (V_n) is compared to infectious virus in uninhibited cultures (V_o) at various dilutions of antisera. The neutralization titer of

the sera is then interpolated as that sera dilution which yields one log reduction in infectious titer (i.e., V_n/V_o = Briefly, 4-fold dilutions of virus (laboratoryadapted and primary isolates) are prepared to yield 5 infectious doses of 0.1 to 100 TCID50 (Tissue Culture Infection Dose) in 20 ul. Serial 3-fold dilutions of sera are also prepared and 20 ul of each serum dilution are incubated with each dilution of virus in duplicate for 60 minutes at room temperature in a 96-well microtiter plate. 10 20 ul of AA5 cells (PHA stimulated PBMCs for primary HIV-1 isolates) are then added to the serum/virus mixtures. Cells are cultured for 7 days by the addition of fresh medium every other day. On the seventh day, supernatant from each well is removed and tested for the presence of reverse 15 transcriptase (RT). Infection in each well is then scored as either positive or negative based on the RT counts, and the infectious dose of virus in each treatment group is calculated using the Reed and Muench (28) formula. neutralization titers represent the reciprocal serum 20 dilution required to reduced infectious dose of virus by one The above culture time is for the prototypic HIV-1LAI isolate tested on the AA5 cell line. In the case of primary isolates, the termination date is usually 11-14 days. Culture conditions for PBMCs is not as demanding since 25 doubling time is restricted. In the case of PBMCs, one day PHA stimulations are used at a final concentration of 1.5 X $10^6/\text{ml}$ on day 0. Half that number of fresh PBMCs are then added again on days 4 and 8. This multiple addition of PBMCs is meant to amplify virus output upon successful 30 infection so that the readout RT signal is strong. Again, the final readout titer for the primary isolate/PBMC is the reciprocal serum dilution which reduces infectious titer by one log.

16. Passive hyperimmune therapy.

Non-HIV-1-infected humans are immunized with the mutant HIV1 gp120 envelope glycoprotein antigens according to a
protocol similar to that described above in section 12. For
5 passive hyperimmune therapy in HIV-1-infected individuals,
blood plasma is taken from mutant HIV-1 gp120 envelope
glycoprotein immunized, non-HIV-1-infected human donors
whose plasma has high levels of neutralizing antibodies.
The plasma is pooled from several donors, purified to remove
10 nonimmunoglobulin proteins and is then sterilized to kill
any other viruses or pathogens. The treated plasma is then
injected into individuals infected with HIV-1, with repeated
injections every week, every two weeks, or every month.

Results

Eukaryotic expression vectors designed to express high levels of HIV-1 $_{\text{LAI}}$ gp120 and HIV-1 $_{\text{IR-FL}}$ gp120 were constructed. 5 The CMV MIE promoter/enhancer was used to drive the transcription of a gene fusion consisting of the human tPA signal sequence fused to mature gp120 (Figures 2 and 7). The complete sequence of the transcription unit from the Hinc II site of the CMV promoter/enhancer to the Not I site just 3' from the stop codon in gp120 is shown in figure 3. This vector was used to transfect COSM5 cells in a transient assay. The transfected cells were labeled with 35-cysteine and the media immunoprecipitated with a CD4-immunoglobulin-Protein A-Sepharose complex. The precipitated products were 15 analyzed using a reducing 10% SDS-PAGE autoradiography (Figure 4). A 120 kD band was detected when PPI4-tPA-gp120_{LAI} was used to transfect COS cells (lane 3). A band migrating with a slightly lower molecular mass was detected when PPI4-tPA-gp120 $_{\mbox{\scriptsize R-FL}}$ was used to transfect COS 20 cells (lane 4). No radiolabeled products were detected in the mock infected cells. Using a sheep polyclonal antibody directed against the highly conserved C-terminal end of HIV-1 gp120 in an ELISA assay, the level of expression of HIV-1 gp120 was determined to be 2350 ng/ml.

25

The PPI4-tPA-gp120_{LAI} vector was then used to stably transfect the dhfr CHO cell line DXB11. Two days post-transfection, the cells were plated at low density in nucleoside-free medium. Eight days post-transfection, surviving clones were isolated and expanded. Individual primary transfectants were tested for gp120 expression using the ELISA method described in the methods section. Several primary CHO transfectants expressed significant quantities (10-120 ng/ml) of gp120 (Figure 5). Three of the highest

expressing clones were then subjected to increasing concentrations of methotrexate in order to amplify, in tandem, the copy number of the dhfr and gp120 genes. Cell lines were established that express high levels of gp120 with rates of secretion greater than 1 mg/liter. These were then used to purify gp120 to homogeneity.

Six CHO cell lines were established, using the procedures described in the methods sections, that express high levels of the following proteins: HIV-1 gp120_{LAI}, gp120_{LAI}-V3⁽⁺⁾, gp120_{LAI}-V3⁽⁺⁾, gp120_{LAI}-V3⁽⁺⁾, and gp120_{IR-FL}-V3⁽⁺⁾. CD4⁽⁺⁾. Metabolic labeling of these cells with ³⁵S-cysteine followed by immunoprecipitation with the human monoclonal antibody F105 and analyzed by SDS-PAGE and autoradiography showed the presence of the gp120 proteins in the culture supernatant (Figure 14). From these cell lines the gp120 proteins were purified to homogeneity. Analysis by SDS-PAGE followed by silver-staining showed the purity of these proteins to be greater than 90% (Figure 15).

20

25

It was shown by FACScan analysis that the two CD4 binding mutants HIV-1gp120 $_{\rm LAI}$ -V3 $^{(\cdot)}$ -CD4 $^{(\cdot)}$ and HIV-1 gp120 $_{\rm IR-FL}$ -V3 $^{(\cdot)}$ -CD4 $^{(\cdot)}$ had no appreciable binding to recombinant cell lines designed to express high levels of human CD4 on their membrane surface (Figure 16, panel 4 and data not shown, respectively).

Discussion

The advantage of using the mutant HIV-1 gp120 envelope 5 glycoproteins as immunogens is that these proteins will not elicit an immune response against the V3 loop, a highly immunodominant epitope on gp120. This is significant because the V3 loop may skew the humoral immune response away from discontinuous epitopes in the CD4-binding site. Mutant HIV-1 10 gp120 envelope glycoproteins having partial and total V3 loop deletions have been made (30). Deletion of the V3 loop therefore exposes the CD4-binding site to the immune system, allowing the immune system to mount a response against this critical region (18). Another advantage of using the mutant 15 HIV-1 gp120 envelope glycoprotein as an immunogen is that it has significantly reduced affinity for cell surface CD4. An efficient humoral immune response depends on the binding of antigen to B cell surface immunoglobulin. The presence of the high-affinity CD4 receptor on large numbers of cells in the body may significantly diminish the ability of native gp120 to induce an effective humoral immune response. The rationale of mutating gp120 at the CD4 binding site is to redirect the mutant HIV-1 gp120 envelope glycoprotein away from cell surface CD4 toward immunoglobulin-bearing B cells, thereby allowing the immune system to mount a response against, inter alia, the CD4-binding site.

References

1. Klatzmann, D.R., et. al. (1990) Immunodeficiency Reviews 2, 43-66.

5

- 2. Lasky, L.A., et. al. (1987) Cell 50, 975-985.
- 3. Maddon, P.J., et. al. (1986) Cell 47, 333-348.
- 10 4. Maddon, P.J., et. al. (1988) Cell 54, 865-874.
 - 5. Maddon, P.J., et. al. (1985) Cell 42, 93-104.
- 6. Maddon, P.J., et. al. (1987) Proc. Natl. Acad. Sci. U.S.A. 84, 9155-9159.
 - Richardson, N.E., et. al. (1988) Proc. Natl. Acad. Sci. U.S.A. 85, 6102-6106.
- 20 8. Chao, B.H., et. al. (1989) J. Biol. Chem. 264, 5812-5817.
 - 9. Arthos, J., et. al. (1989) Cell 57, 469-481.
- 25 10. Wang, J., et. al. (1990) Nature 348, 411-418.
 - 11. Ryu, S.-E., et. al. (1990) Nature 348, 419-426.
- 12. Leonard, C.K., et. al. (1990) J. Biol. Chem. 265, 30 10373-10382.
 - Earl, P.L., et. al. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 648-652.
- 35 14. Helseth, E., et. al. (1991) J. Virol. 65, 2119-2123.

WO 94/22477 PCT/US94/03282

52

- 15. Bolognesi, D.P. (1990) TIBTech 8, 40-45.
- 16. Olshevsky, U., et. al. (1990) J. Virol. 64, 5701-5707.
- 5 17. Steimer, K.S., et. al. (1991) AIDS 5, S135-143.
 - 18. Wyatt, R., et. al. (1992) J. Virol. 66, 6997-7004.
- 19. Zolla-Pazner, S., et. al. (1992) Sem. in Virology 3,203-211.
 - 20. Steimer, K.S., et. al. (1991) Science 254, 105-108.
 - 21. Pollard, S.R., et. al. (1992) EMBO J. 11, 585-591.

15

- 22. Okayama, H. (1983) Mol. Cell. Biol. 3, 280-289.
- 23. Maniatis, T., et. al. (1990) Molecular Cloning, Vol. 1-3.

20

- 24. Thomsen, D.R., et. al. (1984) Proc. Natl. Acad. Sci.
 U.S.A. 81, 659-663.
- 25. Pennica, D., et. al. (1983) Nature 301, 214-221.

25

- 26. Wain-Hobson, S., et. al. (1985) Cell 40, 9-17.
- 27. Koyanagi, Y. (1987) Science 236, 819-822.
- 30 28. Reed, L.J. (1938) Am. J. Hyg., 27, 493-497.
 - 29. Cohn, B.J. et al., (1944) J. Clin. Invest. 23, 417-432.
- 30. Shiow-Her, C., et al. (1992) J. of Cellular Biochem., 35 Supplement 16E, Abstrtact Q105.

SEQUENCE LISTING

- (1) GENERAL INFORMATION:
 - (i) APPLICANT: Progenics Pharmaceuticals, Inc.
 - (ii) TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES THEREOF
 - (iii) NUMBER OF SEQUENCES: 29

 - (iv) CORRESPONDENCE ADDRESS:
 (A) ADDRESSEE: Cooper & Dunham
 - (B) STREET: 30 Rockefeller Plaza
 - (C) CITY: New York
 - (D) STATE: New York
 - (E) COUNTRY: USA
 - (F) ZIP: 10112
 - (V) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk

 - (B) COMPUTER: 1BM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: Patentin Release #1.24
 - (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:

 - (vii) PRIOR APPLICATION DATA:
 (A) APPLICATION NUMBER: US 08/037,816
 (B) FILING DATE: 26-MAR-1993
- (viii) ATTORNEY/AGENT INFORMATION:

 - (A) NAME: White, John P. (B) REGISTRATION NUMBER: 28,678
 - (C) REFERENCE/DOCKET NUMBER: 41190-A-PCT/JPW/AJM
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (212) 977-9550 (B) TELEFAX: (212) 664-0525 (C) TELEX: 422523 COOPUI
- (2) INFORMATION FOR SEQ ID NO:1:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 amino acids
 - (B) TYPE: amino acid (C) STRANDEDNESS: single.
 - (D) TOPOLOGY: Linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
 - Xaa Xaa Xaa Cys Xaa Ile Xaa Xaa Xaa Xaa Xaa Trp Xaa Xaa Xaa
 - Xaa Xaa Ala Xaa Tyr Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa

Ser Xaa Xaa Thr Gly Xaa Xaa Xaa Arg Xaa Gly Xaa

54

35

40

45

- (2) INFORMATION FOR SEG ID NO:2:
 - (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: Linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Thr Leu Pro Cys Arg Ile Lys Gln Phe Ile Asn Met Trp Gln Glu Val

Gly Lys Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser

Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly 35 40 45

- (2) INFORMATION FOR SEQ ID NO:3:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 amino acids (B) TYPE: amino acid

 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: Linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Thr Leu Pro Cys Arg Ile Lys Gin Ile Ile Asn Met Trp Gin Giu Vai 1 10 15

Gly Lys Ala Net Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys Ser 20 25 30

Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly 40

- (2) INFORMATION FOR SEQ ID NO:4:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single (D) TOPOLOGY: Linear
 - (ii) MOLECULE TYPE: DNA (genomic)
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GATCCTGCAG TCACCGTCCT TGACACGATG GATGCAATGA AGAGA

45

- (2) INFORMATION FOR SEQ ID NO:5:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid

(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:	
AAGTETICIC CICGGTETIG TETTITTAAC ACCEAG	36
(2) INFORMATION FOR SEQ ID NO:6:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 36 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6: TTCAGAAGAG GAGCCAGAAC AGAAAAATTG TGGGTC	36
(2) INFORMATION FOR SEQ ID NO:7:	30
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 39 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: DNA (genomic)	
(11) MOLECULE TIPE: DRA (genualic)	
(xi) SEQUENCE DESCRIPTION: SEQ 1D NO:7:	
GGAAAAAGC GGCCGCTCAT TTTTCTCTCT GCACCACTC	39
(2) INFORMATION FOR SEQ ID NO:8:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 33 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:	
GATCGGCGCC AGAGTAGAAA AGTTGTGGGT CAC	33
(2) INFORMATION FOR SEQ ID NO:9:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 49 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	

(fi) MOLECULE TYPE: DNA (genomic)

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
  CTGTAGAAAT TAATTGTACA GGTGCTGGAC ATTGTAACAT TAGTAGAGC
                                                                             49
  (2) INFORMATION FOR SEQ ID NO:10:
       (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 29 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: Linear
     (ii) MOLECULE TYPE: DNA (genomic)
     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
 CTCGAGCATG CATTCGAAGC TCGCTGATC
                                                                            29
 (2) INFORMATION FOR SEQ ID NO:11:
      (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 29 base pairs
           (B) TYPE: nucleic acid
           (C) STRANDEDNESS: single
           (D) TOPOLOGY: linear
     (ii) MOLECULE TYPE: DNA (genomic)
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:
CAATTTATAA ACATGGTGCA GGAAGTAGG
                                                                            29
(2) INFORMATION FOR SEQ ID NO:12:
     (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 29 base pairs
           (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single (D) TOPOLOGY: timear
    (ii) MOLECULE TYPE: DNA (genomic)
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:
CAAATTATAA ACATGGTGCA GGAAGTAGG
(2) INFORMATION FOR SEQ ID NO:13:
     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 3125 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
          (D) TOPOLOGY: Linear
   (ii) MOLECULE TYPE: DNA (genomic)
   (ix) FEATURE:
          (A) NAME/KEY: CDS
         (B) LOCATION: 1555..3115
         (D) OTHER INFORMATION:
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TIGACATTGA TTATTGACTA GTTATTAATA GTAATCAATT ACGGGGTCAT TAGTTCATAG	60
CCCATATATG GAGTTCCGCG TTACATAACT TACGGTAAAT GGCCCGCCTG GCTGACCGCC	120
CAACGACCCC CGCCCATTGA CGTCAATAAT GACGTATGTT CCCATAGTAA CGCCAATAGG	180
GACTITICCAT TGACGICAAT GGGTGGACTA TITACGGTAA ACTGCCCACT TGGCAGTACA	240
TCAAGTGTAT CATATGECAA GTACGCCCCC TATTGACGTC AATGACGGTA AATGGCCCGC	300
CTGGCATTAT GCCCAGTACA TGACCTTATG GGACTTTCCT ACTTGGCAGT ACATCTACGT	360
ATTAGTCATC GCTATTACCA TGGTGATGCG GTTTTGGCAG TACATCAATG GGCGTGGATA	420
GCGGTTTGAC TCACGGGGAT TTCCAAGTCT CCACCCCATT GACGTCAATG GGAGTTTGTT	480
TTGGCACCAA AATCAACGGG ACTTTCCAAA ATGTCGTAAC AACTCCGCCC CATTGACGCA	540
AATGGGCGGT AGGCGTGTAC GGTGGGAGGT CTATATAAGC AGAGCTCGTT TAGTGAACCG	600
TCAGATCGCC TGGAGACGCC ATCCACGCTG TTTTGACCTC CATAGAAGAC ACCGGGACCG	660
ATCCAGCCTC CGCGGCCGGG AACGGTGCAT TGGAACGCGG ATTCCCCGTG CCAAGAGTGA	720
CGTAAGTACC GCCTATAGAC TCTATAGGCA CACCCCTTTG GCTCTTATGC ATGCTATACT	780
GTTTTTGGCT TGGGCCAACA CCCCGTCCTA GATAGGTGAT GGTATAGCTT AGCCTATAGG	840
TGTGGGTTAT TGACCATTAT TGACCACTCC CCTATTGGTG ACGATACTTT CCATTACTAA	900
TCCATAACAT GGCCGCTCTT TGCCACAACT ATCTCTATTG GCTATATGCC AATACTCTGT	960
CCTTCAGAGA CTGACACGGA CTCTGTATTT TTACAGGATG GGGTCCCATT TATTATTTAC	1020
AMATTCACAT ATACAACAAC GCCGTCCCCC GTGCCCGCAG TTTTTATTAA CATGCGGGAT	1080
CTCCACGCGA ATCTCGGGTA CGTGTTCCGG ACATGGGCTC TTCTCCGGTA GCGGCGGAGC	1140
TCCACATCEG AGCETGTECC ATGCCCATGC CTCCAGCGGC TCATGGTEGC TCGGCAGCTC	1200
CTTGCTCCTA ACAGTGGAGG CCAGACTTAG GCACAGGACA ATGCCCACCA CCACCAGTGT	1260
GCCGCACAAG GCCGTGGCGG TAGGGTATGT GTCTGAAAAT GAGCTCGGAG ATTGGGCTCG	1320
CACCGCTGAC GCAGATGGAA GACTTAAGGC AGCGGCAGAA GAAGATGCAG GCAGCTGAGT	1380
TGTTGTATTC TGTAGAGTTG GAGGTAACTC CCGTTGCGGT GCTGTTAACG GTGGAGGGCA	1440
GTGTAGTCTG AGCAGTACTC GTTGCTGCCG CGCGCCCAC CAGACATAAT AGCTGACAGA	1500
CTAACAGACT GTTCCTTTCC ATGGGTCTTT TCTGCAGTCA CCGTCCTTGA CACG ATG Met 1	1557
GAT GCA ATG AAG AGA GGG ETC TGC TGT GTG CTG CTG CTG TGT GGA GCA Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Cys Gly Ala 5 10 15	1605
GTC TTC GTT TCG CCC AGC CAG GAA ATC CAT GCC CGA TTC AGA AGA GGC Val Phe Val Ser Pro Ser Gin Glu 1le His Ala Arg Phe Arg Gly 20 25 30	1653
GCC AGA ACA GAA AAA TTG TGG GTC ACA GTC TAT TAT GGG GTA CCT GTG Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val 35 40 45	1701
TGG AAG GAA GCA ACC ACC ACT CTA TTT TGT GCA TCA GAT GCT AAA GCA	1749

7 /	р L	ys (lu	Ala	Th	r 1h 5:	r Th 5	r Le	u Ph	e Cy	s Ala	a Se	r As	p Al	a Ly	s Ala 65	
TA	IT G	AT A	CA hr	GAG	GT/ Val	L His	T AA' S AS	T GT n Va	T TG L Tr	G GC P Ala 7	e Thr	A CA'	T GCI	C TG	T GT s Va 8	A CCC l Pro O	1797
AC Th	A G/	AC C	cc	AAC Asn 85	Pro	CA/	A GA	4 GT.	A GT. L Va 91	t Lei	G GTA	AA' Asi	r GTC	AC	r Gl	A AAT u Asn	1845
TT Ph	T AA e As	in M	TG et 00	TGG Trp	Lys	AA1 Asr	GA(Mer 10	t Va	A GAJ	CAG Glm	ATC Met	CAT His	Gli	GA'	T ATA	1893
AT Il	C AC e Se 11	r L	TA Bu	TGG Trp	GAT Asp	CAA Glr	AGC Ser 120	. Fer	A AAG	G CCA B Pro	TGT Cys	GTA Val 125	Lys	Leu	ACC The	CCA Pro	1941
CT Let	и Су	T G	IT .	AGT Ser	TTA Leu	Lys 135	Cys	ACT Thr	GA1 Asp	t TTG	GGG Gly 140	Asn	GCT Ala	ACT Thr	AA1 Asr	ACC Thr 145	1989
AA'	T AG	T AC	iT i	AAT Asn	ACC Thr 150	Asn	AGT Ser	AGT Ser	AGC Ser	GGG Gly 155	Glu	ATG Met	ATG Met	ATG Met	GAG GLL 160	Lys	2037
GE	N GA ∕ GL	G A1	e 1	AAA Lys 165	AAC Asn	TGC Cys	TCT Ser	TTC Phe	AAT ASN 170	Ile	AGC Ser	ACA Thr	AGC Ser	ATA Ile 175	Arg	GGT	2085
Lys	GT: Va	G CA I GI 18	n i	MA .ys	GAA Glu	TAT	GCA Ala	TTT Phe 185	Phe	TAT	AAA Lys	CTT Leu	GAT Asp 190	ATA Ile	ATA	CCA Pro	2133
ATA	GAT ASI 195) As	T C	AT ISP	ACT Thr	ACC Thr	AGC Ser 200	TAT	ACG Thr	TTG Leu	ACA Thr	AGT Ser 205	TGT Cys	AAC Asn	ACC Thr	TCA Ser	2181
GTC Val 210	H	AC Th	A C	AG iln	GCC Ala	TGT Cys 215	CCA Pro	AAG Lys	GTA Val	TCC Ser	TTT Phe 220	GAG Glu	CCA Pro	ATT	CCC Pro	ATA Ile 225	2229
CAT	TAT	TG Cy:	T G	la	CCG Pro 230	GCT Ala	GGT Gly	TTT Phe	GCG Ala	ATT Ile 235	CTA Leu	AAA Lys	TGT Cys	AAT Asn	AAT Asn 240	AAG Lys	2277
ACG Thr	Phe	AA Asi	n G	GA / ly 1 45	ACA Thr	GGA Gly	CCA Pro	TGT Cys	ACA Thr 250	AAT Asn	GTC Val	AGC Ser	ACA Thr	GTA Val 255	CAA Gln	TGT Cys	2325
ACA Thr	CAT His	GG/ GL) 260	/ I	TT /	AGG Arg	CCA Pro	GTA Val	GTA Val 265	TCA Ser	ACT Thr	CAA Gin	CTG L eu	CTG Leu 270	TTG Leu	AAT -Asn	GGC	2373
\GT Ser	CTA Leu 275	GC/ Ala	G	AA (iAA ilu	Glu	GTA Val 280	GTA Val	ATT Ile	AGA Arg	Ser	GCC Ala 285	AAT Asn	TTC Phe	ACA Thr	GAC Asp	2421
AT ISD 190	GCT Ala	Lys	A TI	cc A	le	ATA Ile 295	GTA Val	CAG Gln	CTG Leu	AAC Asn	CAA Gln : 300	TCT Ser	GTA Val	GAA Glu	Ile	AAT Asn 305	24 69
GT YS	ACA Thr	AGA	Pr	ro A	ISD I	AAC . Asn .	AAT . Asn	ACA Thr	Arg	Lys 315	AGT /	ATC le	CGT . Arg	Ile	CAG Gln 320	AGG Argʻ	2517
										GGA :							2565

			325				330	ı			335	,		
			Cys		AGT Ser		Ala				Thr		AAA Lys	2613
		Ala			AGA Arg 360	Glu				Asn				2661
	Phe				GGA Gly									2709
				Glu	TTT Phe								Phe	2757
					AGT Ser									2805
			Ser		ATC Ile									2853
					GTA Val 440									2901
	Gly				TCA Ser									2949
					AAC Asn									2 99 7
					AAT Asn									3045
					TTA Leu									3093
Arg	GTG Val	Vat			AAA Lys	T GA	GCGG	CCGC	:					3125

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 520 amino acids (B) TYPE: amino acid (D) TOPOLOGY: Linear
- (11) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10 15

Ata Val Phe Vat Ser Pro Ser Gin Glu Ite His Ata Arg Phe Arg Arg 20 25 30

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro

														6	0	
		:	35					4	0				4	5		
۷a	l Tr 5	p L	/s G	lu .	Ala	Th	r Th 5	r Th 5	r Le	u Phe	Cys	Ala 60	s Se	r Ası	o Ali	B Ly:
Al 5	a Ty 5	r As	sp T	hr I	Glu	Va 1	l Hi D	s Ası	n Va	l Trp	Ala 75	Thr	· His	s Ala	E Cy	s Va 80
Pr	o Th	r As	p P	ro i	Asn 85	Pro	o Gla	n Glu	ı Va	l Val	Leu	ı Vəl	. Ası	n Vai	Thi	- Glu
Asi	n Ph	e As	in Mi	et 1	Trp	Lys	S Ası	n Asp	105	t Val	Glu	Gln	Me1	110	Glu	Asp
11	e Il	e Se 11	r Lo 5	eu 1	Γrp	Asp	Gli	Ser 120	· Lei	ı Lys	Pro	Cys	Val 125	Lys	Leu	. Thr
Pr	130	u Cy D	's Va	al S	Ser	Leu	135	Cys	Thr	Asp	Leu	Gly 140		Ala	Thr	Asr
Th:		n Se	r Se	er A	lsn	Thr 150	Asr	ser	Ser	Ser	Gly 155	Glu	Het	Het	Net	Glu 160
Lys	s Gly	/ Gl	u 11	le L	.ys 165	Asn	Cys	Ser	Phe	170	ile	Ser	Thr	Ser	I le 175	Arg
Gly	/ Lys	s Va	l GI 18	In L 30	.ys	Glu	Tyr	Ala	Phe 185	Phe	Tyr	Lys	Leu	Asp 190	Ile	ile
Pro	Ile	19	p As 5	in A	sp	Thr	Thr	Ser 200	Туг	The	Leu	Thr	Ser 205	Cys	Asn	Thr
Ser	Val 210	11	e Th	r G	ln	Ala	Cys 215	Pro	Lys	Val	Ser	Phe 220	Glu	Pro	Ile	Pro
11e 225	His	Ty	г Су	'S A	la	Pro 230	Ala	Gly	Phe	Ala	I l e 235	Leu	Lys	Cys	Asn	Asn 240
Lys	Thr	Pho	AS -	n G	ly 45	Thr	Gly	Pro	Cys	Thr 250	Asn	Val	Ser	Thr	Val 255	Gln
Cys	Thr	His	G G L	y 1 0	le	Arg	Pro	Val	Val 265	Ser	Thr	Gln	Leu	Leu 270	Leu	Asn
Gly	Ser	Let 275	ı Al	a G	lu	Glu	Glu	Val 280	Val	Ile	Arg	Ser	Ala 285	Asn	Phe	Thr
Asp	Asn 290	Ala	Ly:	s Ti	hr	Ile	11e 295	Val	Gln	Leu	Asn	Gln 300	Ser	Val	Glu	Ile
Asn 305	Cys	Thr	Arı	g Pi	ro :	Asn 310	Asn	Asn	Thr	Arg	Lys 315	Ser	Il⊕	Arg	Ile	Gln 320
Arg	Gly	Pro	GLY	7 A1	rg /	Ala	Phe	Val	Thr	11e 330	Gly	Lys	Ite	Gly	A s n 335	Het
Arg	Gln	Ala	H11 340) ()	/B /	Asn	Ite	Ser	Arg 345	Ala	Lys	Trp	Asn	Ala 350	Thr	Leu
Lys	Gln	1 le 355	Ala	S S c	r I	Lys	Leu	Arg 360	Glu	Gln	Phe	Gly	Asn 365	Asn	Lys	Thr
Ile	1 l e 370	Phe	Lys	Gl	n S	Ser	Ser 375	Gly	Gly	Asp	Pro	Glu 380	Ile	Val	Thr	His
Ser	Phe	Asn	Cvs	. GI	v 6	ilv	Glu	Phe	Phe	Tvr	Cve	Agn	SAF	The	CLO	

Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp Ser Thr Glu Gly Ser Asn

٠				405					41	0				4	15			
As	n Thi	r Glu	420	Ser	Asp	Thr	ile	2 Th 42	r Le 5	u Pr	o Cy	/s Ar	g (le L [.] 30	ys	Gln	•	
Ph	e ilo	435	Met	Trp	Gln	Glu	۷a (440	. GL	y Ly	s Al	a Me	t Ty 44		a Pi	0	Pro		
H	Ser 450	Gly	Gln	ile	Arg	Cys 455	Ser	Se	r As	n Il	e Th 46	r Gl	y Le	u L	eu	Leu		
Th:	Arg	Asp	Gly	Gly	Asn 470	Asn	Asn	ASI	n Gl	y Se 47	r Gl 5	u Il	e Ph	e Ar		Pro 480		
Gly	/ Gly	Gly	Asp	Met 485	Arg	Asp	Asn	Trp	49	g Sei	r Gl	u Le	и Ту	r Ly 49		Tyr		
			Lys 500					Gl y 505	/ Vai	l Ala	∌ Pr	o Th	r Ly 51		ai	Lys		
Arg	Arg	Val 515	Val	Gln	Arg	Glu	Lys 520									-		
(2)	INF	ORMA	TION	FOR	SEO	ID N	10:1	5:										
		() () ()	RUENC A) LE B) TY C) ST C) TO	NGTH PE: RAND POLO	: 15 nucl EDNE GY:	32 b eic ss: line	acid sing	pai d gle										
			.ECUL		PE:	DNA	(ger	וחסר	c}									
		(A) NAI	ME/KI	ON:	11		:										
	(xi)	SEQ	UENCI	E DES	SCR []	PTIO	N: S	EQ :	1D N	0:15	:							
ITG let 1	GAT Asp	GCA Ala	ATG / Met l	MG A	lGA (GGG (CTC Leu	TGC Cys	TGT Cys 10	GTG Val	CTG Leu	CTG Leu	Let	TG1 Cys	G	GA ly		48
iCA N l a	GTC Val	TTC Phe	GTT 1 Val S 20	CG C	CC A	AGC (CAG Gln	GAA Glu 25	ATC Ile	CAT His	GCC Ala	CGA Arg	Phe 30	Arg	A A	GA rg		96
igc	GGC	AGA (Arg 1 35	GTA G Val G	iAA A ilu L	AG 1 YB L	ITG 1	ree rp 40	GTC Vel	ACA Thr	GTC Val	TAT	TAT Tyr 45	GGG	GTA Val	Pi	CT PO	14	44
TG al	TGG . Trp 50	AAA (Lys (GAA G Glu A	CA A	hr T	CC / hr 1 55	hr i	CTA Leu	TTT Phe	TGT Cys	GCA Ala 60	TCA Ser	GAT Asp	GCT	A)	VA 78	19	92
CA la 65	TAT (SAT / Asp 1	ICA G	lu V	TA C al H 70	AT A	AT (GTT Val	TGG Trp	GCC Ala 75	ACÀ Thr	CAT His	GCC Ala	TGT Cys	V۹	ra il 30	24	40
CC (ACA (AC C	CC A	AC CI sn Pi 85	CA C ro G	AA G in G	AA (STA Val	GTA Val 90	TTG Leu	GAA Glu	AAT Asn	GTA Val	ACA Thr 95	GA Gl	u u	28	38
AT 1	TTT A	IAC A	TG T	GG A	M A	AT A	AC A	ITG	GTA	GAA	CAG	ATG	CAG	GAG	GA	Ť	33	36

AT.	A AT	c AG e Se 11:	r Lei	A TGI	GA1	r CAA	AGO Ser 120	Let	A AAC	G CCA	TG1	7 GT/ 5 Val 125	Ly	A TT. S Le	A ACC u Thr	384
		u Cy:					Cys					n Ala			T ACC	
ACT Thi 145	- Asi	r GA1	T AGO	GAC	GGA Gly 150	, Thr	Met	GAG Glu	AGA Arg	GGA Gly 155	Glu	ATA Ile	AAA Lys	A AAC S ASI	TGC Cys 160	
TC1 Sei	Phe	AA1	T ATO	165	Thr	AGC Ser	ATA	AGA Arg	GAT Asp 170	Glu	GTG Val	GLr	AAA Lys	GA4 Glu 175	TAT Tyr	528
				. Lys					Pro					ASF	ACC Thr	576
			Leu					Thr					Gln		TGT Cys	624
		Ite										Cys			GCT Ala	672
	Phe					Cys									GGA Gly 240	720
Pro	TGT Cys	Lys	AAT	GTC Val 245	AGC Ser	ACA Thr	GTA Val	CAA Gln	TGT Cys 250	ACA Thr	CAT His	GGA Gly	ATT	AGG Arg 255	Pro	768
GTA Val	GTA Val	TCA Ser	ACT Thr 260	CAA Gin	CTG Leu	CTG Leu	CTA Leu	AAT Asn 265	GGC Gly	AGT Ser	CTA Leu	GCA Ala	GAA Glu 270	GAA Glu	GAG	816
GTA Val	GTA Val	Ile 275	AGA Arg	TCT	GAC Asp	AAT Asn	TTC Phe 280	ACG Thr	AAC Asn	AAT Asn	GCT Ala	AAA Lys 285	ACC Thr	ATA	ATA Ile	864
GTA Val	CAG Gln 290	CTG Leu	AAA Lys	GAA Glu	TCT Ser	GTA Val 295	GAA Glu	ATT Ile	AAT Asn	TGT Cys	ACA Thr 300	AGA Arg	CCC Pro	ASD	AAC Asn	912
Asn 305	Thr	Arg	Lys	Ser	11e 310	His'	ile	GGA Gly	Pro	Gly 315	Arg	Ala	Phe	Tyr	Thr 320	960
The	Gly	Glu	He	11e 325	Gly	Asp	:le		Gln 330	Ala	His	Cys	Asn	1 l e 335	Ser	1008
Arg	Ala	Lys	Тгр 340	Asn	Asp	Thr	Leu	Lys 345	Gln	ile	Val	He	Lys 350	Leu	Arg	1056
GAA Glu	CAA Gln	TTT Phe 355	GAG Glu	AAT Asn	AAA Lys	Thr	ATA 1 l e 360	GTC Val	TTT Phe	AAT Asn	His	TCC Ser 365	TCA Ser	GGA Gly	GGG Gly	1104
GAC Asp	CCA Pro 370	GAA Glu	ATT Ile	GTA Val	Met	CAC His 375	AGT Ser	TTT Phe	AAT Asn	Cys	GGA Gly 380	GGA Gly	GAA Glu	TTT Phe	TTC Phe	1152
TAC	TGT	AAT	TCA	ACA	CAA	CTG	TTT	AAT	AGT	ACT	TEE	AAT	AAT	AAT	ACT	1200

1 y r 385	Cys	Asr	Ser	Thr	390	Leu	Phe	Asn	Ser	1hr 395	Trp	Asn	Asn	Asn	thr 400	
GAA	GGG	TCA Ser	AAT Asn	AAC Asn 405	ACT Thr	GAA	GGA Gly	AAT Asn	ACT Thr 410	He	ACA Thr	CTC	CCA Pro	TGC Cys 415	AGA Arg	1248
ATA	AAA Lys	CAA Gln	ATT lle 420	ATA Ile	AAC Asn	ATG Met	TGG Trp	CAG Gln 425	GAA Glu	GTA Val	GGA Gly	AAA Lys	GCA Ala 430	ATG Met	TAT Tyr	1296
GCC Ala	CCT Pro	CCC Pro 435	ATC 11e	AGA Arg	GGA Gly	CAA Gln	ATT Ile 440	AGA Arg	TGT Cys	TCA Ser	TCA Ser	AAT Asn 445	ATT	ACA Thr	GGG Gly	1344
			ACA Thr													1392
TTC Phe 465	AGA Arg	CCT Pro	GGA Gly	GGA Gly	GGA Gly 470	GAT Asp	ATG Met	AGG Arg	GAC Asp	AAT Asn 475	TGG Trp	AGA Arg	AGT Ser	GAA Glu	TTA Leu 480	1440
TAT Tyr	AAA Lys	TAT Tyr	AAA Lys	GTA Val 485	GTA Val	AAA Lys	ATT Ile	Glu	CCA Pro 490	TTA Leu	GGA Gly	GTA Val	Ala	CCC Pro 495	ACC Thr	1488
AAG Lys	GCA Ala	AAG Lys	AGA Arg 500	AGA Arg	GTG Val	GTG Val	Gln	AGA Arg 505	GAA Glu	AAA Lys	T GA	GCGG	cccc	:		1532

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 507 amino acids (8) TYPE: amino acid

 - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ale Arg Phe Arg Arg 20 2530

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ale Thr Thr Thr Leu Phe Cys Ale Ser Asp Ale Lys 50 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Glu Asn Val .Thr Glu 85 90 '95

His Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met Gln Glu Asp 100 105 110

Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Thr Leu Asn Cys Lys Asp Val Asn Ala Thr Asn Thr 130 135 140

Thr 145	Asn	Asp	Ser	Glu	Gly 150	Thr	Met	Glu	Arg	Gly 155	Glu	ile	Lys	Asn	Cys 160
Ser	Phe	Asn	He	Thr 165	Thr	Ser	lle	Arg	Asp 170	Glu	Val	Gln	Lys	Glu 175	tyr
Ala	Leu	Phe	Tyr 180	Lys	Leu	Asp	Val	Val 185	Pro	lle	Asp	Asn	Asn 190	Asn	Thr
Ser	Tyr	Arg 195	Leu	ile	Ser	Cys	Asp 200	Thr	Ser	Val	lle	Thr 205	Gln	Ala	Cys
Pro	Lys 210	Ile	Ser	Phe	Glu	Pro 215	Ile	Pro	Ile	His	Tyr 220	Cys	Ala	Pro	Ala
Gly 225	Phe	Ala	lle	Leu	Lys 230	Cys	Asn	Asp	Lys	Thr 235	Phe	Asn	Gly	Lys	Gly 240
Pro	Cys	Lys	Asn	Val 245	Ser	Thr	Val	Gln	Cys 250	Thr	His	Gly	Ile	Arg 255	Pro
Val	Val	Ser	Thr 260	Gln	Leu	Leu	Leu	Asn 265	Gly	Ser	Leu	Ala	Glu 270	Glu	Glu
Val	Val	1 le 275	Arg	Ser	Asp	Asn	Phe 280	Thr	Asn	Asn	Ala	Lys 285	Thr	He	Ile
Val	Gln 290	Leu	Lys	Glu	Ser	Val 295	Glu	ile	Asn	Cys	Thr 300	Arg	Pro	Asn	Asn
Asn 305	Thr	Arg	Lys	Ser	I l e 310	His	He	Gly	Pro	Gly 315	Arg	Ala	Phe	Туг	Thr 320
Thr	Gly	Glu	Ile	I le 325	Gly	Asp	Ile	Arg	Gln 330	Ala	His	Cys	Asn	11e 335	Ser
Arg	Ala	Lys	1rp 340	Asn	Asp	Thr	Leu	Lys 345	Gln	:le	Val	Ile	Lys 350	Leu	Arg
Glu	Gln	Phe 355	Glu	Asn	Lys	Thr	1 l e 360	Val	Phe	Asn	His	Ser 365	Ser	Gly	Gly
Agn															
nsp	Pro 370	Glu	Ile	Val	Met	His 375	Ser	Phe	Asn	Cys	Gly 380	Gly	Glu	Phe	Phe
•	370				Met Gln 390	375				•	380	•			
1yr 385	370 Cys	Asn	Ser	Thr	Gln	375 Leu	Phe	Asn	Ser	Thr 395	380 Trp	Asn	Asn	Asn	Thr 400
Tyr 385 Glu	370 Cys Gly	Asn Ser	Ser Asn	Thr Asn 405	Gln 390	375 Leu Glu	Phe Gly	Asn Asn	Ser Thr 410	Thr 395	380 Trp Thr	Asn	Asn Pro	Asn Cys 415	Thr 400 Arg
Tyr 385 Glu Ile	370 Cys Gly Lys Pro	Asn Ser Gln Pro 435	Ser Asn Ile 420 Ile	Thr Asn 405 Ile Arg	Gin 390 Thr Asn Gly	375 Leu Glu Het Gin	Phe Gly Trp Ile 440	Asn Asn Gln 425 Arg	Ser Thr 410 Glu Cys	Thr 395 Ile Val	380 Trp Thr Gly Ser	Asn Leu Lys Asn 445	Asn Pro Ala 430 Ile	Asn Cys 415 Het	Thr 400 Arg Tyr
Tyr 385 Glu Ile Ala	370 Cys Gly Lys Pro Leu 450	Asn Ser Gln Pro 435 Leu	Ser Asn Ile 420 Ile	Thr Asn 405 Ile Arg	Gln 390 Thr Asn Gly	375 Leu Glu Het Gln Gly 455	Phe Gly Trp 11e 440 Gly	Asn Gln 425 Arg	Thr 410 Glu Cys	Thr 395 Ile Val Ser	380 Trp Thr Gly Ser Asn 460	Asn Leu Lys Asn 445 Gly	Asn Pro Ala 430 Ile	Asn Cys 415 Het Thr	Thr 400 Arg Tyr Gly
Tyr 385 Glu Ile Ala Leu Phe 465	370 Cys Gly Lys Pro Leu 450 Arg	Asn Ser Gin Pro 435 Leu	Ser Asn Ile 420 Ile Thr	Thr Asn 405 Ile Arg Gly	Gln 390 Thr Asn Gly Asp Gly 470	375 Leu Glu Met Gln Gly 455 Asp	Phe Gly Trp Ile 440 Gly	Asn Asn Gln 425 Arg Ile	Ser Thr 410 Glu Cys Asn	Thr 395 Ile Val Ser Glu Asn 475	380 Trp Thr Gly Ser Asn 460 Trp	Asn Leu Lys Asn 445 Gly	Asn Pro Ala 430 Ile Thr	Asn Cys 415 Het Thr Glu	Thr 400 Arg Tyr Gly Ile Leu 480
Tyr 385 Glu Ile Ala Leu Phe 465	370 Cys Gly Lys Pro Leu 450 Arg	Asn Ser Gln Pro 435 Leu Pro	Ser Asn Ile 420 Ile Thr Gly Lys	Thr Asn 405 Ile Arg Gly Val 485	Gln 390 Thr Asn Gly	375 Leu Glu Met Gln Gly 455 Asp	Phe Gly Trp Ile 440 Gly Met	Asn Asn Gln 425 Arg Ile Arg	Thr 410 Glu Cys Asn Asp	Thr 395 Ile Val Ser Glu Asn 475 Leu	380 Trp Thr Gly Ser Asn 460 Trp	Asn Leu Lys Asn 445 Gly	Asn Pro Ala 430 Ile Thr	Asn Cys 415 Het Thr Glu	Thr 400 Arg Tyr Gly Ile Leu 480

,		ruk	25.A	10	RU: 17;	
	(i) SEQUENC	E CH	ADAC	TE	167166	_

- SEQUENCE CHARACTERISTICS:

 (A) LENGTH: 1484 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
- (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1..1474
 - (D) OTHER INFORMATION:
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AT(G GA C As	T GC. P Ali	A ATI	t Ly	G AG/ B Arg	A GGC G Gly	CTC	; TGC	TG1 Cys	Val	CTC Let	CTC	i CT(I Lel	3 TG1 Cy1 15	T GGA B Gly	48
GCA Ala	GTI Va	C TTI L Phi	C GT1 e Vai 20	Ser	CCC Pro	AGC Ser	CAG Gln	GAA Glu 25	Ile	CA1	GCC	CGA Arg	170 Phe 30	Arg	AGA Arg	96
GGC	GC(C AG/ B Arg 35	3 Thr	GAA	Lys	tTG Leu	TGG Trp 40	Val	ACA Thr	GTC	TAT	TAT Tyr 45	GGG	GTA Val	CCT Pro	144
GTG Val	TGC Trp 50) Lys	GAA Glu	GCA Ala	ACC	ACC Thr 55	ACT	CTA Leu	TTT Phe	TGT Cys	GCA Ala 60	Ser	GAT Asp	GCT Ala	Lys	192
GCA Ala 65	TAT	GAT Asp	ACA Thr	GAG	GTA Val 70	His	AAT Asn	GTT Val	TGG Trp	GCC Ala 75	ACA Thr	CAT His	GCC Ala	TGT Cys	GTA Val 80	240
CCC Pro	ACA Thr	GAC Asp	CCC Pro	AAC Asn 85	CCA Pro	CAA Gln	GAA Glu	GTA Val	GTA Val 90	TTG Leu	GTA Val	AAT Asn	GTG Val	ACA Thr 95	GAA Glu	288
AAT Asn	TTT Phe	AAC Asn	ATG Met 100	TGG Trp	AAA Lys	AAT Asn	GAC Asp	ATG Met 105	GTA Val	GAA Glu	CAG Gln	ATG Met	CAT His 110	GAG Glu	GAT Asp	336
ATA 1le	ATC Ile	AGT Ser 115	TTA Leu	TGG Trp	GAT Asp	CAA Gin	AGC Ser 120	CTA Leu	AAG Lys	CCA Pro	TGT Cys	GTA Val 125	AAA Lys	TTA Leu	ACC Thr	384
PFO	CTC Leu 130	TGT Cys	GTT Val	AGT Ser	TTA Leu	AAG Lys 135	TGC Cys	ACT Thr	GAT Asp	TTG L e u	GGG Gly 140	AAT Asn	GCT Ala	ACT Thr	AAT Asn	432
ACC Thr 145	AAT Asn	AGT Ser	AGT Ser	AAT Asn	ACC Thr 150	AAT Asn	AGT . Ser	AGT Ser	Ser	GGG Gly 155	GAA Glu	ATG Met	ATG Met	Met	GAG Glu 160	480

AAA GGA GAG ATA AAA AAC TGC TCT TTC AAT ATC AGC ACA AGC ATA AGA Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg 165 170 175

GGT AAG GTG CAG AAA GAA TAT GCA TTT TTT TAT AAA CTT GAT ATA ATA Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile 180 185 190

CCA ATA GAT AAT GAT ACT ACC AGC TAT ACG TTG ACA AGT TGT AAC ACC Pro Ite Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr 195 200 205

576

TC Se	A GT r Va 21	ı II	T AC e Th	A CA	G GC n Al	C TG: a Cy: 21:	s Pr	A AAI O Ly:	G GT/ s Val	A TCC	Pho 220	e Gli	Pro	TA A	T CCC e Pro	672
AT. 11 22	e Hi	T TA	T TG	T GC s Al	c cca a Pro 23:	o Ala	r GG1 B Gly	T TT1	r GCC	ATT 11e 235	: Le	A AAA	TG1 Cys	T AA' S Asi	T AAT n Asn 240	720
AA(G AC s Th	G TT r Ph	C AA e Asi	T GG/ n Gl: 24!	y Thi	A GGA r Gly	CC#	TG1 Cys	ACA Thr 250	· Asn	GTC Val	: AGC	ACA Thr	Val 255	A CAA L Gln	768
TG1 Cys	T AC	A CA r Hi	T GG/ S Gl: 26	y Ile	AGO Arg	G CCA 9 Pro	GTA Val	Val 265	Ser	ACT Thr	GLA	CTG Leu	Leu 270	Leu	AAT Asn	816
GGC	AG' Sei	CT.	u Ali	A GAA	GAA Glu	A GAG	GTA Val 280	Val	ATT	AGA Arg	TCT Ser	GCC Ala 285	Asn	TTC Phe	ACA Thr	864
GAC Asp	AA1 Asr 290	n Ali	AA/	A ACC	ATA Ile	ATA Ile 295	Val	CAG Gin	CTG Leu	AAC Asn	CAA Gln 300	Ser	GTA Val	GAA	ATT Ile	912
AAT Asn 305	Cys	AC/	GG1 Gly	GCT Ala	GGA Gly 310	CAT His	TGT Cys	AAC Asn	ATT	AGT Ser 315	AGA Arg	GCA Ala	AAA Lys	TGG Trp	AAT Asn 320	960
GCC	ACT Thr	TT/	Lys	CAG Gln 325	He	GCT	AGC Ser	AAA Lys	TTA Leu 330	AGA Arg	GAA Glu	CAA Gln	TTT Phe	GGA Gly 335	Asn	1008
AAT Asn	Lys	ACA	ATA Ile 340	lle	TTT Phe	AAG Lys	CAA Gin	TCC Ser 345	TCA Ser	GGA Gly	GGG Gly	GAC Asp	CCA Pro 350	GAA Glu	ATT	1056
GTA Val	ACG Thr	CAC His 355	Ser	TTT Phe	AAT Asn	TGT Cys	GGA Gly 360	GGG Gly	GAA Glu	TTT Phe	TTC Phe	TAC Tyr 365	TGT Cys	AAT Asn	TCA Ser	1104
ACA Thr	CAA Gln 370	CTG	TTT Phe	AAT Asn	AGT Ser	ACT Thr 375	TGG Trp	TTT Phe	AAT Asn	Ser	ACT Thr 380	TGG Trp	AGT Ser	ACT Thr	GAA Glu	1152
GGG Gly 385	TCA Ser	AAT Asn	AAC Asn	ACT Thr	GAA Glu 390	GGA Gly	AGT Ser	GAC Asp	Thr	ATC 1le 395	ACA Thr	CTC Leu	CCA Pro	TGC Cys	AGA Arg 400	1200
ATA Ile	AAA Lys	CAA Gln	TTT Phe	ATA Ile 405	AAC Asn	ATG Het	TGG Trp	Gln	GAA Glu 410	GTA Val	GGA Gly	AAA Lys	Ala	ATG Met 415	Tyr	1248
GCC Ala	CCT Pro	CCC Pro	ATC Ile 420	AGC Ser	GGA Gly	CAA Gln	Il●	AGA Arg 425	TGT Cys	TCA Ser	TCA Ser	Asn	ATT Ile 430	ACA Thr	GGG	1296
CTG L e u	CTA Leu	TTA Leu 435	ACA Thr	AGA Arg	GAT Asp	GGT	GGT Gly 440	AAT Asn	AAC Asn	AAC . Asn .	Asn	GGG Gly 445	TCC Ser	GAG Glu	ATC Ile	1344
he	AGA Arg 450	CCT Pro	GGA Gly	GGA Gly	Gly	GAT ASP 455	ATG Met	AGG Arg	GAC A	Asn '	TGG Trp 460	AGA . Arg :	AGT Ser	GAA Glu	TTA Leu	1392
TAT Tyr 165	AAA Lys	TAT Tyr	AAA Lys	Val	GTA Val 470	AAA /	ATT Ile	GAA Glu	Pro I	TTA (Leu (475	GGA Gly	GTA (Val	GCA (Pro	ACC Thr 480	1440
MG	GCA	AAG	AGA	AGA	GTG	GTG (CAG .	AGA (GAA /	w 1	T GAI	GCGG	CGC			1484

Lys Ala Lys Arg Arg Val Val Gin Arg Glu Lys
485

- (2) INFORMATION FOR SEG ID NO:18:
 - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 491 amino acids

 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

Pro Thr Asp Pro Asn Pro Gin Giu Val Val Leu Val Asn Val Thr Giu 85 90 95

Asn Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110

Tie Tie Ser Leu Trp Asp Gin Ser Leu Lys Pro Cys Vai Lys Leu Thr 115 120 125

Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn 130 135 140

Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Met Glu 145 150 160

Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg 165 170 175

Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile 180 185 190

Pro Ile Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr 195 200 205

Ser Val Ite Thr Gin Ala Cys Pro Lys Val Ser Phe Glu Pro Ite Pro 210 215

Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asn 225 230 235 240

Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln 245 250 255

Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Asn 260 265 270

Gly Ser Leu Ala Glu Glu Glu Val Val 11e Arg Ser Ala Asn Phe Thr 275 280 280

													•	•		
Ası	290	Ala	Lys	Thr	Ile	1 l e 295	Val	Glr	Leu	Asr	Glr 300	Ser	· Va	l Gt	ı Ile	
Asr 305	n Cys	Thr	Gly	Ala	Gly 310	His	Cys	Asn	lle	Ser 315	Arg	Ala	Lys	i Tr	Asn 320	
Ala	1hr	Leu	Lys	Gln 325	Ile	Ala	Ser	Lys	Leu 330	Arg	Gtu	Glr	Phe	GL 335	/ Asn	
Asr	1 Lys	Thr	I le 340		Phe	Lys	Gln	Ser 345		Gly	Gly	Asp	9rc 350		ille	
Val	Thr	His 355	Ser	Phe	Asn	Cys	Gly 360		Glu	Phe	Phe	Tyr 365		Asr	Ser	
Thr	Gln 370	Leu	Phe	Asn	Ser	Thr 375	Trp	Phe	Asn	Ser	Thr 380	Trp	Ser	Thr	Glu	
Gly 385	Ser	Asn	Asn	Thr	Glu 390	Gly	Ser	Asp	Thr	1 l e 395	Thr	Leu	Pro	Cys	Arg 400	
Ile	Lys	Gln	Phe	11e 405	Asn	Met	Trp	Gln	Glu 410	Val	Gly	Lys	Ala	Met 415		
Ala	Pro	Pro	1 l e 420	Ser	Gly	Gln	Ile	Arg 425	Cys	Ser	Ser	Asn	1 le 430		Gly	
Leu	Leu	Leu 435	Thr	Arg	Asp	Gly	Gly 440	Asn	Asn	Asn	Asn	Gly 445	Ser	Glu	Ile	
Phe	Arg 450	Pro	Gly	Gly	Gly	Asp 455	Het	Arg	Asp	Asn	1rp 460	Arg	Ser	Glu	Leu	
Tyr 465	Lys	Tyr	Lys	Val	Val 470	Lys	Ile	Glu	Pro	Leu 475	Gly	Val	Ala	Pro	Thr 480	
Lys	Ala	Lys	Arg	Arg 485	Val	Vai	Gln	Arg	Glu 490	Lys						
(2)	INFO	RMAT	ION	FOR	SEQ	ID N	O:19):	•							
	(i)	SEQ	UENC	E CH	ARAC	TERI	STIC	:2:								
	•••	ÇA.) LE	NGTH	: 14	48 b	-	pair	8							
					nuc l EDNE											
					GY:									•		
	(ii)	MOL	ECUL	E TY	PE:	DNA	(gen	omic	:)							
	(ix)	FEA	TURE	:												
					EY: ON:		470									
					INFO											
	(xi)	SEQ	JENC	E DE	SCRI	PTIO	N: S	EQ 1	D NO	:19:						
ATG	GAT	GCA /	ATG /	AAG .	AGA (GGG	CTC	TGC	TGT	STG	CTG	cts	CTG	TGT	CGA	48
Met 1	Asp	Ala I	let	Lys 5	Arg (Gly	Leu	Cys	Cys 10	Val	Leu	Leu	Leu	Cys 15	Gly	~~
	GTC T													Arg		96

GGC GGC AGA GTA GAA AAG TTG TGG GTC ACA GTC TAT TAT GGG GTA CCT Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

A TCA GAT GCT AAA 197 9 Ser Asp Ala Lys -)	GCA T Ala S 60	TGT Cys	TTT Phe	CTA	r Thi	C AC	A AC	M GC	A GA	D LA	TG(Tr; S(Val
CAT GCC TGT GTA 240 His Ala Cys Val 80	ACA CA	GCC Ala 75	TGG	GTT Val	T AA1 S ASI	l Hi	G GT U Va 7	A GA	T AC p Th	GA As	Iy	GCA Ala 65
AAT GTA ACA GAA 288 Asn Val Thr Glu 95	GAA A/ Glu As	TTG Leu	GTA Val 90	GTA Val	GAA	A CAA	n Pr	C AA O As 8	C CC P Pr	GA(ACA Thr	CCC
ATG CAG GAG GAT 336 Met Gln Glu Asp 110	CAG A1 Gln Me	GAA Glu	GTA Val	ATG Met 105	AAC ASN	A AAT S Asi	G AA. P Ly:	t Iul	C AT n Me 10	AA(TTT Phe	CAT
GTA AAA TTA ACC 384 Val Lys Leu Thr 125	Cys Va	CCA Pro	AAG Lys	Leu	AGC Ser 120	r CAA	G GA' P Asi	A TGO	r Le	Ser 115	ATC	ATA Ile
GCT ACT AAT ACC 432 Ala Thr Asn Thr	AT GC Isn Al I40	Val.	GAT Asp	AAG Lys	Cys	AAT J Asn 135	r IT/	T ACT	GT1	Cys	CTC Leu 130	CCA Pro
ATA AAA AAC TGC 480 Ile Lys Asn Cys 160	AA AT ilu Il	GGA (Gly (Arg	GAG Glu	ATG Het	Thr	GG/ Gly 150	GAG Glu	AGC Ser	GAT Asp	AAT Asn	ACT Thr 145
CAG AAA GAA TAT 528 Gln Lys Glu Tyr 175	al Gl	Glu 1	170	Arg	ile	Ser	Thr	165	ile	ASN	rne	ser
AAT AAT ACC 576 Asn Asn Asn Thr 190	sp Asi	ile /	Pro	185	Val	Asp	Leu	Lys	180	rne	reu	A La
ACA CAG GCC TGT 624 Thr Gln Ala Cys 205	le Thi 205	/al I	Ser '	Thr	Asp 200	Cys	Ser	ille	Leu	Arg 195	Tyr	ser
TGT GCC CCG GCT 672 Cys Ala Pro Ale	yr Cys 20	lis T	ile i	Pro	Ite	215	Glu	Phe	Ser	11e	210	750
AAT GGA AAA GGA 720 Asn Gly Lys Gly 240	he Asr	hr P 35	Lys 1	ASP	ASI	Cys	230	LEU	116	ALA	ne .	25
GGA ATT AGG CCA 768 Gly Ile Arg Pro 255	is Gly	hr H	250	Gin	Val	Thr	Ser	Val 245	Asn	Lys	ys i	ro (
GCA GAA GAA GAG 816 Ala Glu Glu Glu 270	nu Ala	er L	ily s	Nan (265	Leu	Leu	Leu	GLN	260	Ser	81 3	8(\
	a Lys 285	sn A	lsn A	Thr /	Phe ' 280	ASN	Asp	Ser	Arg	275	at 1	al V
GGT GCT GGA CAT 912 Sly Ala Gly His	r Gly O	ys Ti 30	sn C	ile A	Glu I	Val 295	Ser	Glu	Lys	.eu i	in i 90	al G 2
AA CAG ATA GTT 960 ys Gln Ile Val 320	A AAA U Lys	CT TI hr Le 15	sp T	AT G	rgg /	AAA '	GCA Ala 310	Arg .	AGT Ser	le :	sn i	GT A VS A DS
						+++ 4	~ 4 4 ~		404	TA 4		7 A A

116	e Lys	Let	1 AFG	325		Phe	Glu	Asn	330 F A a		' !le	· Val	Phe	335	His	
TCC Ser	TCA Ser	GGA	GGG Gly 340	Asp	CCA Pro	GAA Glu	ATT	GTA Val 345	Met	His	AGT Ser	Phe	AAT Asn 350	Cys	GGA Gly	1056
			Phe			AAT Asn							Ser			1104
		Asn				TCA Ser 375						Asn				1152
	Pro					CAA Gln										1200
						CCC Pro										1248
						TTA Leu										1296
						CCT Pro		Gly								1344
						TAT Tyr 455										13 9 2
				Lys		AAG Lys			Val						TG	1439
GCC	GCCG	C														1448

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 479 amino acids
 (B) TYPE: amino acid

 - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Het Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro $35 \hspace{1cm} 40 \hspace{1cm} 45$

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Glu Asn Val Thr Glu

														′	1	
					8	5				9	0				9	5
Hi	s Ph	e A	lsn	10:	t Tr D	p Ly	s As	n As	n Me 10	t Va 5	l Gle	u Gli	n Me	t Gli	n Gl	u Asp
п	e Il	e .9	er 15	Lei	u Tr	p As	p Gl	n Sei 120	r Le O	u Ly	s Pro	Cy:	125		s Le	u Thr
Pr	o Le 13		ys	Va	l Th	rle	u Asi 13:		s Ly	s As	o Val	Asr 140		a Thi	r Ası	n Thr
7h:	r As 5	n A	sp	Ser	- Gl	u Gl: 15	y Thi	r Mei	c Gli	u Ar	g Gly 155	Gli	ı Ile	Lys	s Ası	160
Sei	r Ph	e A	sn	He	16!	r Thi	r Sei	r Ile	e Arg	9 Asp 170	Glu)	ı Val	Glr	Lys	Gl:	ı Tyr
Ala	a Le	u P	he	Tyr 180		s Lei	ı Asp	Val	Val 185		ile	Asp	Asn	190		1hr
Ser	Ty	T A	rg 95	Leu	H	e Ser	· Cys	Asp 200	Thr	Ser	Val	Ile	1hr 205	Gln	Ala	Cys
	211)					215	i				220				Ala
Gl y 225	Pho	e A	la	lie	Leu	230	Cys	Asn	Asp	Lys	Thr 235	Phe	Asn	Gly	Lys	Gly 240
Pro	Cy:	i Ly	ys	Asn	Val 245	Ser	Thr	Val	Gln	Cys 250	Thr	His	Gly	Ile	Arg 255	Pro
Val	Val	Se		Thr 260	Gln	leu	Leu	Leu	Asn 265	Gly	Ser	Leu	Ala	Glu 270	Glu	Glu
Val	Val	1 l 27	.e '5	Arg	Ser	Asp	Asn	Phe 280	Thr	Asn	Asn	Ala	Lys 285	Thr	Ile	Ile
Val	290	l				•	295				Cys	300				
Cys 305	Asn	11	e :	Ser	Arg	Ala 310	Lys	îrp	Asn	Asp	Thr 315	Leu	Lys	Gln	Ile	Val 320
Ile	Lys	Le	U	Arg	Gl u 325	Gln	Phe	Glu	Asn	Lys 330	Thr	He	Val	Phe	Asn 335	His
Ser	Ser	Gl	y (Gly 340	Asp	Pro	Glu	ile ,	Val 345	Het	His	Ser	Phe	Asn 350	Cys	Gly
		35	>					360			Leu		365		Thr	Trp
Asn	Asn 370	Ası	n 1	[hr	Glu	Gly	Ser 375	Asn	Asn	Thr	Glu	380	Asn	Thr	Ile	Thr
.eu 585	Pro	CY	8 /	\rg	Ile	Lys 390	Gln	Ile	Ile	Asn	Het 395	Trp	Gln	Glu	Val	Gly 400
.ys	Ala	Met	1	Уľ	Ala 405	Pro	Pro	Ile	Arg	Gly 410	Gln	Ile .	Arg		Ser 415	Ser
Isn	Ile	Thr	- G	ly 20	Leu	Leu	Leu	Thr	Arg 425	Asp	Gly	Gly		Asn 430	Glu	Asn

Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly

450

va 46	l Al S	a Pr	o Th	r Ly	'S Al 47	a Ly	s Ar	g Ar	g Va	ι Va 47		n Ar	g Gl	u Ly	5	
(2) IN	FORM	ATIO	N FO	R SE	Q ID	NO:	21:								
	(CHAR.										•	
					TH: : nu				irs							
					NDEDI											
					LOGY											
	(i	i) M	OLEC	ULE	TYPE	: DN	A (g	enom	ic)							
	(i		EATU													
					/KEY: TION:			4								
					RINI											
	(xi) SI	EQUE	NCE (DESCR	RIPTI	ON:	SEQ	ID A	10:21	1:					
ATG	GAT	GC	A ATI	G AAI	G AGA	GGC	CTO	: TG(: TGI	GTO	. cto	CTO	CTO	TGI	GGA	48
Met	: Asp	AL	Me 1	t Lys	s Arg	Gly	Lei	Cys	Cys	: Val	Lei	Leu	Leu	Cys	GLY	
1				:	•				10)				15	i	
GCA	GTC	TTO	GT1	TC	CCC	AGC	CAC	GAJ	ATC	CAT	GCC	CGA	TTC	AGA	AGA	96
Ala	Val	Phe	• Val	Ser)	· Pro	Ser	Glr	ı Glı 25	ı Ile	His	: Ala	Arg	Phe 30		Arg	
GGC	GCC	AGA	ACA	GAA		TTE	tee	: ctr	. 454	GTC		TAT		CT.	CCT	4//
Gly	Ala	Arg	; Thr	Gli	Lys	Leu	Trp	Val	Thr	Val	Tyr	Tyr	Gly	Val	Pro	144
		35	i				40)				45				
GTG	TGG	AAG	GAA	GCA	ACC	ACC	ACT	CTA	TTT	TGT	GCA	TCA	GAT	GCT	AAA	192
val	1 rp	Lys	GLU	Ala	Thr	Thr 55	Thr	Leu	Phe	Cys	Ala 60		Asp	Ala	Lys	
GCA	TAT	CAT		C40	-											
Ala	Tyr	ASD	The	Glu	GTA Val	His	ASI	Val	Tro	GCC	ACA	CAT	GCC	TGT	GTA	240
65		•			70					75				-,-	80	
CCC	ACA	GAC	CCC	AAC	CCA	CAA	GAA	GTA	GTA	TTG	GTA	AAT	GTG	ACA	GAA	288
Pro	Thr	Asp	Pro	Asn 85	Pro	Gln	Glu	Val	Val 90	Leu	Val	Asn	Val	Thr 95	Glu	
	•••															
AAT	Phe	AAC	ATG	TGG	AAA Lys	AAT	GAC	ATG	GTA	GAA	CAG	ATG	CAT	GAG	GAT	336
	• •••	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	100	,	-,-	Agii	vab	105	*40.	410	utn	HEL	110	Glu	ASP	
ATA	ATC	AGT	TTA	TGG	GAT	CAA	AGC	CTA	AAG	CCA	TGT	GTA	AAA	TTA	ACC	384
Ile	He	Ser	Leu	Trp	Asp	Gln	Ser	Leu	Lys	Pro	Cys	Val	Lys	Leu	Thr	304
		115					120					125				
CCA	CTC	TGT	GTT	AGT	TTA	AAG	TGC	ACT	GAT	TTG	GGG	AAT	GCT	ACT	AAT	432
FIG	130	Lys	Val	Ser	Fen	135	Cys	IRF	ASP	Leu	140	ASN	ALB	Thr	Asn	
ACC	AAT	AGT	ACT	AAT	ACC	AAT	ACT	ACT	ACC	ccc		470				/80
Thr	Asn	Ser	Ser	Asn	Thr	Asn	Ser	Ser	Ser	Glv	Glu	Met	Net	Het	Glu	480
145					150					155		•			160	
AAA	GGA	GAG	ATA	**	AAC	TGC	TCT	TTC	AAT	ATC	AGC	ACA	AGC	ATA	AGA	528
Lys	Gly	Glu	He	Lys 165	Asn	Cys	Ser	Phe	Asn 170	He	Ser	Thr	Ser		Arg	
														175		
GGT Gly	AAG	GTG Val	CAG	AAA L v=	GAA	TAT	GCA	TTT Ph=	TTT	TAT	AAA	CTT	GAT	ATA	ATA	576
,	-,,			-73		. ,,		L 116	7 11 5		- 73	-	vob		116	

					180)					185	;					19	0			
C: P:	CA CO	ATA	A:	SP /	AAT Asn	GA As	T A(T A	hr S	GC 1 er 1 00	ГАТ	AC Th	G TI	G A	hr :	igt Ser	Су	T A/ S A:	AC sn	ACC Thr	624
T (• [GTC Val 210	11	e i	ACA Thr	CA Gl	G GC	a C	GT C ys P 15	CA A	AG .ys	GT/ Va	A TC	r Pi	TT (ne (AG Lu	CC. Pr	1A A	11 .e	CCC Pro	. 677
A1 1 U 22	e I	CAT His	TA Ty	T I	GT ys	GC:	C CC a Pr 23	O AI	CT G	GT T ly P	TT he	GCC Ala	3 AT 3 I U 23	e Le	TA A	AA YS	TG'	T AA B AS	iT in	AAT Asn 240	720
AA Ly	G /	ACG Thr	TT Ph	C A	AT Sn	GG/ GL: 24!	y Th	A GO	A CO	CA T	GT ys	ACA Thr 250	Ası	T GT n Va	C A	GC er	AC/ Thi	V GT Va 25	ι	CAA Gln	768
TG Cy	s I	ACA Thr	CA Hi	SG	GA ly 60	ATI	AG Ar	G CC g Pr	A G1 O Va	it v	TA a l 65	TCA Ser	AC'	T CA	A C	TG eu	CT (Let 270	Le	G	AAT Asn	816
GG Gl	C A	igt Ser	CT. Le 27:	U A	CA la	GA/ Glu	GA,	A GA J Gl	G GT u Va 28	il Vi	TA	ATT	AG	t TC	r A	CC la 85	AAT Asn	TT Ph	C	ACA Thr	864
GA:	PA	AT ISN 190	GC AL	T A	AA ys	ACC Thr	AT/	A AT.	A GT • Va 5	A CA	AG ln	CTG Leu	AAC	CA Gli 30	n Si	er er	GTA Val	GA/ Gli	4	ATT Ile	912
AA1 Asi 305	1 6	GT YS	AC/ Thi	GI GI	it ly	GCT Ala	GG/ GLy 310	/ H1:	T TG S Cy	T AJ	IC In	ATT Ile	AGT Ser 315	· Arı	A GO	A	AAA Lys	TGG) /	MT Nsn 520	960
GCC	A TI	CT hr	TT#	L	/S	CAG Gln 325	ATA	GC1	r AG	C AA	' S	TTA Leu 330	AGA Arg	GL	A CA	n i	TTT Phe	GG/ Gl y 335	, ,	WT Isn	1008
AAT Asn	L	AA ys	ACA Th <i>r</i>	A1 11 34	•	ATC	TTT Phe	Lyt	G CA	TC Se 34	•	TCA Ser	GGA Gly	GGG	GA / As	P	CCA Pro 350	GAA Glu	1	ite	1056
GTA Val	AC 4T	ודו	CAC His 555	AG Se	T T	TTT Phe	AAT Asn	TG1 Cys	GG/ GL) 360	/ Gl	G (GAA Glu	TTT Phe	TTC Phe	74 79 36	r(rgt Cys	AAT Asn	5	CA	1104
ACA Thr	GI 37	n l	eu.	TT Ph	T A	LAT LSN	AGT Ser	ACT Thr 375	TCC	TT Ph	T /	MAT Asin	AGT Ser	ACT Thr 380	Tr	G /	GT Ger	ACT Thr	G	AA lu	1152
366 31 y 385	Se	A A	LST LST	AA Asi	C A	nr	GAA Glu 390	GGA Gly	AGT Ser	GAI AS	C #	lhr	ATC I le 395	ACA Thr	CT Le	C C	CA	TGC Cys	A	GA rg 00	1200
lta Le	Ly	A C	in	TT: Pho	e I	TA le 05	AAC Asn	ATG Met	GTG Val	Gli	, (ilu ilu ilo	GTA Val	GGA Gly	Ly	A G	la	ATG Met 415	T.	AT yr	1248
icc	Pr	T C o P	CC	AT0 116 420	3	GC er	GGA Gly	CAA Gln	ATT	AGA Arg 425	9 C	GT :	TCA Ser	TCA Ser	AA1 Asi	۱ ۱	TT le 30	ACA Thr	G	GG Ly	1296
TG C U	CT/ Lei	n F	TA eu 35	ACA Thr	A	GA :	GAT Asp	GGT Gly	GGT Gly 440	Asn	A	AC A	AAC Asn	AAT Asn	GG(GL) 445	, S	CC	GAG Glų	A1	rc i.e	1344
ne	AGA Arg	3 P	CT FO	GGA Gly	G	GA (Gly	GAT Asp	ATG Met	AGG Arg	G	AC /	MT Nsn	TGG Trp	AGA	A A	GT (GAA Glu	T1	A N	1392

TAT AAA TAT AAA GTA GTA AAA ATT GAA CCA TTA GGA GTA GCA CCC ACC Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr 465 470 475 480 AAG GCA AAG AGA AG AGTGGTGCAG AGAGAAAAAT GAGCGGCCGC 1484

- (2) INFORMATION FOR SEQ ID NO:22:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 484 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

. Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Val Asn Val Thr Glu 85 90 95

Asn Phe Asn Met Trp Lys Asn Asp Met Vai Glu Gin Met His Glu Asp 100 105 110

Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ale Thr Asn . 130 135 140

Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Met Glu 145 155 160

Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg 165 170 175

Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile 180 185 190

Pro Ile Asp Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr 195 200 205

Ser Val Ite Thr Gin Ala Cys Pro Lys Val Ser Phe Glu Pro Ite Pro 210 215 220

Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asn 225 230 235 240

Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln 245 250 255

Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn

			260)				265	5				270)	
Gly	y Sei	27:	Ala 5	Gli	GLL	Gli	۷a۱ 280		110	e Arq	Ser	Ala 285		n Phe	Thr
Asp	290	n Ala	a Lys	Thr	· Ite	295	val	Glr	n Lei	J Asr	GLr 300		· Val	Gli	ille
Asr 305	Cys	Thr	Gly	/ Ala	Gly 310	His	Cys	ASF	ıle	Ser 319	Arg	Ala	Lys	îrp	Asn 320
Ala	Thr	Lec	l Lys	325	Ile	Ala	Ser	Lys	330		Glu	Gla	Phe	Gly 335	Asn
Asr	Lys	Thr	11e	Ile	Phe	Lys	Gln	Ser 345	Ser	Gly	Gly	Asp	Pro 350		Ile
Val	Thr	His 355	Ser	Phe	Asn	Cys	Gly 360	Gly	Glu	Phe	Phe	7yr 365	Cys	Asn	Ser
Thr	Gln 370	Leu	Phe	Asn	Ser	Thr 375	Trp	Phe	Asn	Ser	Thr 380	Trp	Ser	Thr	Glu
Gly 385	Ser	Asn	Asn	Thr	Glu 390	Gly	Ser	Asp	Thr	I l e 395	Thr	Leu	Pro	Cys	Arg 400
Ile	Lys	Gln	Phe	11e 405	Asn	Met	Val	Gln	Glu 410	Val	Gly	Lys	Ala	Met 415	Tyr
Ala	Pro	Pro	11e 420	Ser	Gly	Gln	Ile	Arg 425	Cys	Ser	Ser	Asn	Ile 430	Thr	Gly
Leu	Leu	Leu 435	Thr	Arg	Asp	Gly	Gly 440	Asn	Asn	Asn	Asn	Gly 445	Ser	Glu	Ile
Phe	Arg 450	Pro	Gly	Gly	Gly	Asp 455	Met	Arg	Asp	Asn	Trp 460	Arg	Ser	Glu	Leu
Tyr 465	Lys	Tyr	Lys	Val	Val 470	Lys	Ile	Glu	Pro	Leu 475	Gly	Val	Ala	Pro	Thr 4 8 0
Lys	Ata	Lys	Arg												

- (2) INFORMATION FOR SEQ ID NO:23:
 - (i) SEGUENCE CHARACTERISTICS: (A) LENGTH: 1448 base pairs (B) TYPE: nucleic acid

 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: Linear
 - (ii) MOLECULE TYPE: DNA (genomic)
 - (ix) FEATURE:

 - (A) NAME/KEY: CDS
 (B) LOCATION: 1..1438
 (D) OTHER INFORMATION:
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:
- ATG GAT GCA ATG AAG AGA GGG CTC TGC TGT GTG CTG CTG CTG TGT GGA Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10 15
- GCA GTC TTC GTT TCG CCC AGC CAG GAA ATC CAT GCC CGA TTC AGA AGA ALa Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

GG	c gg y Gl	y Ar	A G1 g Va 5	A GA	A AA u Ly:	G TT(S Let	3 TG0 3 Trp 40	o Va	C AC/	GTC Val	TA1	T TA	r GL	G GT. y Va	A CCT L Pro	144
GT Va	G TG L Tr	p Ly	A GA	A GC u Al	A ACI	C ACC	Thr	CT/	A TTI J Phe	TG1	GCA Ala 60	Ser	GAT Asj	GC'	T AAA B Lys	192
	a Ty					l His					The				r GTA s Val 80	
					n Pro					Leu					GAA Glu	
CA1	TTI Pho	AA 1 2 As	C AT n Me 10	t Tri	AAA D Lys	AAT ASN	AAC Asn	Met 105	Val	GAA Glu	GLn	ATG Met	GLr 110	GLL	GAT J ASP	336
			r Le					Leu					Lys		ACC Thr	384
Pro	130	Cy:	T GT	r act	Leu	AAT Asn 135	Cys	AAG Lys	GAT Asp	GTG Val	AAT Asn 140	Ala	ACT Thr	AAT Asn	ACC Thr	432
hr 145	Asr	GA1	F AGI	GAC	GGA Gly 150	Thr	ATG Met	GAG Glu	AGA Arg	GGA Gly 155	GAA Glu	ATA	Lys	AAC	TGC Cys 160	480
ct	TTC Phe	AAT ASI	ATC	165	Thr	AGC Ser	ATA Ile	AGA Arg	GAT Asp 170	GAG Glu	GTG Val	CAG Gln	Lys	GAA Glu 175	Tyr	528
la	CTT	TT1 Phe	TA1 Tyr 180	Lys	CTT Leu	GAT Asp	GTA Val	GTA Val 185	CCA Pro	ATA Ile	GAT Asp	AAT Asn	AAT Asn 190	AAT Asn	ACC Thr	576
GC er	TAT	AGG Arg 195	Lec	ATA Ile	AGT Ser	TGT Cys	GAC Asp 200	ACC Thr	TCA Ser	GTC Val	TTA Ile	ACA Thr 205	CAG Gln	GCC Ala	TGT Cys	624
CA ro	AAG Lys 210	Ιle	TCC Ser	Phe	GAG Glu	CCA Pro 215	ATT Ile	CCC Pro	ATA Ile	CAT His	TAT Tyr 220	TGT Cys	GCC Ala	CCG Pro	GCT Ala	672
l y 25	Phe	Ala	Ile	CTA Leu	Lys 230	Сув	Asn	Asp	Lys	Thr 235	Phe	Asn	Gly	Lys	Gly 240	720
ro	Cys	Lys	Asn	GTC Val 245	Ser	Thr	Val	Gln	Cys 250	Thr	His	Gly	He	Arg 255	Pro	768
al	Vel	Ser	7hr 260		Leu	Leu	Leu	A s n 265	Gly	Ser	Leu	Ala	Glu 270	Glu	Glu	816
A	GTA Val	ATT Ile 275	AGA Arg	TCT Ser	GAC Asp	Asn	TTC Phe 280	ACG Thr	AAC Asn	AAT Asn	Ala	AAA Lys 285	ACC Thr	ATA Ile	ATA Ile	864
				GAA Glu						Cys						912
T	AAC	ATT	AGT	AGA	GCA	AAA	TGG .	AAT	GAC	ACT	TTA	AAA	CAG	ATA	GTT	960

305	s Ası	n Il	e Se	r Arı	310	lys	Tr	o Asr	n Asp	315	Leu	J LY	s Gli	n II	e Val 320	
ATA !le	Ly:	A TT	A AG	G GA	ı Glr	ttt Phe	GAC	AAT J Asn	AAA Lys 330	Thr	ATA	GTC Val	Pho	AA 1 Asi 333	T CAC h His	1008
TCC Ser	TC#	GG/	GG(Gl) 34(/ Asp	CCA Pro	GAA	ATT	GTA Val 345	Met	CAC His	AGT Ser	Phe	AA1 Asr 350	Cys	GGA Gly	1056
GGA Gly	GAA Glu	771 Phe 355	Phe	TAC	TGT Cys	AAT Asn	TCA Ser 360	Thr	CAA Gin	CTG Leu	TTT Phe	AAT Asn 365	Ser	AC1	TGG Trp	1104
AAT Asn	AAT Asn 370	Asr	ACT Thr	GAA Glu	GGG	Ser 375	AAT Asn	AAC Asn	ACT Thr	GAA Glu	GGA Gly 380	Asn	ACT	ATC	ACA Thr	1152
CTC Leu 385	CCA Pro	TGC Cys	AGA Arg	ATA	AAA Lys 390	CAA Gln	ATT	ATA Ile	AAC Asn	ATG Met 395	GTG Val	CAG Gln	GAA Glu	GTA Val	GGA Gly 400	1200
AAA Lys	GCA Ala	ATG Met	TAT Tyr	GCC Ala 405	CCT Pro	CCC Pro	ATC ile	AGA Arg	GGA Gty 410	CAA Gln	ATT Ile	AGA Arg	TGT Cys	TCA Ser 415	TCA Ser	1248
AAT Asn	ATT Ile	ACA Thr	GGG Gly 420	CTG Leu	CTA Leu	TTA Leu	ACA Thr	AGA Arg 425	GAT Asp	GGT Gly	GGT Gly	ATT	AAT Asn 430	GAG Glu	AAT Asn	1296
GGG Gly	ACC Thr	GAG Glu 435	ATC Ile	TTC Phe	AGA Arg	Pro	GGA Gly 440	GGA Gly	GGA Gly	GAT Asp	Met	AGG Arg 445	GAC Asp	AAT Asn	TGG Trp	1344
arg	AGT Ser 450	GAA Glu	TTA Leu	TÄT Tyr	Lys	TAT Tyr 455	AAA Lys	GTA Val	GTA Val	Lys	ATT Ile 460	GAA Glu	CCA Pro	TTA Leu	GGA Gly	1392
GTA /al i65	GCA Ala	CCC Pro	ACC Thr	AAG Lys	GCA Ala 470	AAG Lys	AGA Arg	AGA Arg	Val	GTG Val 475	CAA Gln	AGA Arg	GAA Glu	AAA Lys	T	1438
AGC	GGCC	GC														1448

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 479 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys' Val Leu Leu Leu Cys Gly 1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro $35 \ \ \, 40 \ \ \, 45$

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 60

A (a		r Ası	o Thi	r Gli	y Val 70	His	Asn	val	l Trp	75		His	. Ala	Cys	val 80
Pro	The	r Ası	o Pro	8 o 8	n Pro	Glr	Glu	val	Val 90		Glu	AST	Val	Thr 95	Glu
His	Phe	e Asr	100	t Tr;	lys	Asn	Asn	Met 105		Glu	Glm	Met	Gln 110		Asp
Ile	Ile	Ser 115		ı Trp	Asp	Gln	Ser 120		Lys	Pro	Cys	Val 125		Leu	Thr
Pro	130	Cys	val	Thr	Leu	Asn 135	Cys	Lys	Asp	Val	Asn 140		Thr	Asn	Thr
1hr 145		ASF) Ser	Glu	150	Thr	Met	Glu	Arg	Gly 155	Glu	ile	Lys	Asn	Cys 160
Ser	Phe	Asr	ile	165		Ser	lle	Arg	170	Glu	Val	Gln	Lys	Glu 175	Tyr
Ala	Leu	Phe	180		Leu	Asp	Val	Val 185	Pro	ile	Asp	Asn	Asn 190	Asn	Thr
Ser	Туг	Arg 195		i Ite	Ser	Cys	Asp 200	Thr	Ser	Val	Ile	1hr 205	Gln	Ala	Cys
	210				Glu	215					220			Pro	Ala
Gly 225	Phe				Lys 230					235					240
Pro	Cys	Lys	Asn	Val 245	Ser	Thr	Val	Gln	Cys 250	Thr	His	Gly	Ile	Arg 255	Pro
Val	Val	Ser	Thr 260		Leu			265			•		270		
		275		•	Asp	Asn	Phe 280	Thr	Asn	Asn	Ala	Lys 285	Thr	Ile	Ile
	290		Lys	Glu		295	Glu				300	-	Ala	Gly	His
305	Asn		Ser		Ala 310					315				Ile	320
	Lys			325	Gln				330					335	
			340		Pro			345					350	Cys	Gly
		Phe 355			Cys		360					365		Thr	Trp
	370					375					380				Thr
385	Рго	•	Arg		Lys 390					395					Gly 400
	Ala			405	Pro				410					415	Ser
\sn	Ile	Thr	Gly	Leu	Leu	Leu	Thr .	Arg	Asp	Gly	Gly	He	Asn	Glu	Asn

Gly	Thr	Glu 435	Ite	Phe	Arg	Pro	Gly 440	Gly	Gly	Asp	Met	Arg 445	Asp	Asn	Trp
-----	-----	------------	-----	-----	-----	-----	------------	-----	-----	-----	-----	------------	-----	-----	-----

- Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly 450 460
- Val Ala Pro Thr Lys Aia Lys Arg Arg Val Val Gln Arg Glu Lys 465 470 475
- (2) INFORMATION FOR SEQ ID NO:25:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1571 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: DNA (genomic)
 - (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1..1567
 - (D) OTHER INFORMATION:
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AT	G GA	T GC	A AT	G AAG	AG	N GGG	CTC	TGC	: TG1	GTO	CTO	CTG	CTO	TGI	GGA	48
Me	t As _i 1	PAL	a Me	t Lys	Arg	Gly	Leu	Cys	10		Let	Leu	Let	Cys 15	Gly	
GC/ Ala	A GTO	C TT(C GT1 E Val	Ser	Pro	AGC Ser	CAG Gln	GAA Glu 25	ıIle	CAT His	GCC Ala	CGA Arg	TTC Phe 30	Arg	AGA Arg	96
GGC	GCC Ala	AG/ Arg 35	The	GAA Glu	Lys	TTG Leu	TGG Trp 40	Val	ACA Thr	GTC	TAT Tyr	TAT Tyr 45	Gly	GTA Val	CCT Pro	144
GT G Val	TGC Trp 50	Lys	GAA Glu	GCA Ala	ACC	ACC Thr 55	Thr	CTA	TTT Phe	TGT Cys	GCA Ala 60	Ser	GAT Asp	GCT ALa	Lys	192
GCA Ala 65	Tyr	GAT Asp	ACA Thr	GAG	GTA Val 70	CAT His	AAT Asn	GTT Val	TGG Trp	GCC Ala 75	Thr	CAT His	GCC	TGT Cys	GTA Val 80	240
CCC Pro	ACA Thr	GAC Asp	CCC Pro	AAC Asn 85	CCA Pro	CAA Gln	GAA Glu	GTA Val	GTA Val 90	TTG Leu	GTA Val	AAT Asn	GTG Val	ACA Thr 95	GAA Glu	288
AAT Asn	TTT Phe	AAC Asn	ATG Met 100	TGG Trp	AAA Lys	AAT Asn	gac Asp	ATG Met 105	GTA Val	GAA Glu	CAG Gln	ATG Met	CAT His 110	GAG Glu	GAT Asp	336
ATA Ile	ATC	AGT Ser 115	TTA Leu	TGG Trp	GAT Asp	CAA Gln	AGC Ser 120	CTA Leu	AAG Lys	CCA Pro	TGT Cys	GTA Val 125	AAA Lys	TTA Leu	ACC Thr	384
CCA Pro	CTC Leu 130	TGT Cys	GTT Val	AGT Ser	TTA Leu	AAG Lys 135	TGC Cys	ACT Thr	GAT Asp	TTG Leu	GGG Gly 140	AAT Asn	GCT Alm	ACT Thr	AAT Asn	432
						AAT Asn										480
						TGC Cvs										528

				16	5				170)				179	5	
GG	T AA	G GT(G CAC L Glr 180	1 LY	A GAZ	A TAI	GCA Ala	171 Pho 185	Pho	TA1	Lys	CTT Leu	GAT ASE 190	110	ATA A	576
Pro	A AT	A GA1 2 Asp 195) Asr	GA1	T ACT	Thr	AG0 Ser 200	Tyr	ACC Thr	TTC	ACA Thr	AGT Ser 205	Cys	AAC Asr	ACC Thr	624
		llte					Pro					Glu			CCC Pro	672
ATA 11e 225	His	TAT Tyr	TGT Cys	GCC	230) Ala	GGT	TTT Phe	GCG	Ile 235	Leu	Lys	TGT Cys	AAT	AAT Asn 240	720
AAG Lys	Thr	Phe	AAT Asn	GGA Gly 245	Thr	GGA Gly	CCA Pro	TGT Cys	ACA Thr 250	Asn	GTC Val	AGC Ser	ACA Thr	GTA Val 255	Gln	768
												CTG Leu				816
GGC Gly	AGT Ser	CTA Leu 275	Ala	GAA Glu	GAA Glu	GAG Glu	GTA Val 280	Val	ATT	AGA Arg	TCT Ser	GCC Ala 285	AAT Asn	TTC Phe	ACA Thr	864
GAC Asp	AAT Asn 290	Ala	AAA Lys	ACC Thr	ATA Ile	ATA Ile 295	GTA Val	CAG Gln	CTG Leu	AAC Asn	CAA Gln 300	TCT Ser	GTA Val	GAA Glu	ATT	912
AAT Asn 305	TGT Cys	ACA Thr	AGA Arg	CCC Pro	AAC Asn 310	AAC Asn	AAT Asn	ACA Thr	AGA Arg	AAA Lys 315	AGT Ser	ATC Ile	CGT Arg	ATC Ile	CAG Gln 320	960
AGG Arg	GGA Gly	CCA Pro	GGG Gly	AGA Arg 325	GCA Ala	TTT Phe	GTT Val	ACA Thr	ATA Ile 330	GGA Gly	AAA Lys	ATA Ile	GGA Gly	AAT Asn 335	ATG Met	1008
AGA Arg	CAA Gln	GCA Ala	CAT His 340	TGT Cys	AAC Asn	ATT Ile	AGT Ser	AGA Arg 345	GCA Ala	AAA Lys	TGG Trp	AAT Asn	GCC Ala 350	ACT Thr	TTA Leu	1056
AAA Lys	CAG Gln	ATA ile 355	GCT Ala	AGC Ser	AAA Lys	TTA Leu	AGA Arg 360	GAA Glu	CAA Gln	TTT Pho	GGA Gly	AAT Asn 365	AAT Asn	AAA Lys	ACA Thr	1104
Ile	ATC Ile 370	TTT Phe	AAG Lys	CAA Gln	TCC Ser	TCA Ser 375	GGA Gly	GGG Gly	GAC ASP	Pro	GAA Glu 380	ATT	GTA Val	ACG Thr	CAC His	1152
				Gly					Tyr			TCA Ser				1200
TTT Phe	AAT Asn	AGT Ser	Thr	TGG Trp 405	TTT Phe	AAT Asn	AGT Ser	Thr	TGG Trp 410	AGT Ser	ACT Thr	GAA Glu	Gly	TCA Ser 415	AAT Asn	1248
AAC Asn	ACT Thr	Glu	GGA Gly 420	AGT Ser	GAC Asp	ACA Thr	ile	ACA Thr 425	CTC Leu	CCA Pro	TGC Cys	AGA Arg	ATA ile 430	AAA Lys	CAA Gln	1296
TTT . Phe	lle	AAC Asn	ATG Met	GTG Val	CAG Gln	Glu	GTA Val	GGA Gly	AAA Lys	GCA Ala	ATG Met	TAT	GCC Ala	CCT Pro	CCC Pro	1344

		Gly								ATT					TTA Leu	1392
ACA Thr 465	AGA Arg	GAT Asp	GGT Gly	GGT Gly	AAT Asn 470	AAC Asn	AAC Asn	AAT Asn	GGG Gly	TCC Ser 475	GAG Glu	ATC Ile	TTC Phe	AGA Arg	CCT Pro 480	1440
										AGT Ser						1488
										GCA Ala						1536
			GTG Val					TGA	GCG	G CC	:GC					1571

- (2) INFORMATION FOR SEQ ID NO:26:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 522 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: Linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Het Asp Ala Het Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg $20 \hspace{1cm} 25 \hspace{1cm} 30 \hspace{1cm}$

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

Pro Thr Asp Pro Asn Pro Gin Glu Val Val Leu Val Asn Val Thr Glu 85 90 95

Asn Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110

lie lie Ser Leu Trp Asp Gin Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn 130 135 140

Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Het Edu 145 155 160

Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg 165 170 175

Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile 185 190

Pro Ite Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr

		19	>				200)				20:	5		
Se	r Va 21	1 I I	e Th	r Gli	n Ala	B Cy:	s Pro	Ly:	s Va	l Se	220	e Glu	J Pro	lle	Pro
21 to	e Hi: 5	s. Ty	r Cy:	s Ala	230		a Gly	/ Pho	e Ala	235		ı Lys	Cys	Asr	4sn 240
Lys	s Thi	Pho	e Asr	Gly 245		GLY	y Pro	Cys	250		val	Ser	Thr	Val 255	
Cys	. Thr	· His	s Gly 260	/ Ile	Arç) Pro	val	Va (265		Thr	Glm	Leu	270		Asn
Gly	/ Ser	275		Glu	Gli	, Glu	val 280		. Ile	Arg	Ser	Ala 285		Phe	Thr
Asp	290		Lys	Thr	lle	295	Val	Gln	Leu	Asn	Gln 300		Val	GLu	He
Asn 305	Cys	Thr	Arg	Pro	310	Asn	Asn	Thr	Arg	Lys 315		Ile	Arg	He	Gln 320
Arg	Gly	Pro	Gly	Arg 325	Ala	Phe	Val	Thr	1le 330	Gly	Lys	Ile	Gly	Asn 335	Met
Arg	Gin	Ala	His 340		Asn	lle	Ser	Arg 345		.Lys	Trp	Asn	Ala 350	Thr	Leu
Lys	Gln	11e 355		\$er	Lys	Leu	Arg 360	Glu	Gln	Phe	Gly	Asn 365	Asn	Lys	Thr
lle	1 le 370	Phe	Lys	Gln	Ser	Ser 375	Gly	Gly	Asp	Pro	Glu 380	Ile	Val	Thr	His
Ser 385	Phe	Asn	Cys	Gly	Gly 390	Ğlu	Phe	Phe	Tyr	Cys 395	Asn	Ser	Thr	Gln	Leu 400
				405			Ser		410					415	
Asn	Thr	Glu	Gly 420	Ser	Asp	Thr	Ile	Thr 425	Leu	Pro	Cys	Arg	1 le 430	Lys	Gln
		435					Val 440					445			
ile	Ser 450	Gly	Gln	Ile	Arg	Cys 455	Ser	Ser	Asn	Ile	Thr 460	Gly	Leu	Leu	Leu
465					470		Asn			475					480
				485			Asn		490					495	
Lys	Val	Val	Lys 500	Ile	Glu	Pro	Leu	Gly 505	Val	Ala	Pro		Lys 510	Ala	Lys
Arg	Arg	Val 515	Val	Gin	Arg	Glu	Lys 520								

- (2) INFORMATION FOR SEQ ID NO:27:
 - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1532 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genom	si c	:)
--------------------------------	------	-----

- (ix) FEATURE:

 (A) NAME/KEY: CDS

 (B) LOCATION: 1..1522

 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

	(x	i) \$	EQUE	NCE	DESC	RIPT	ION:	SEQ	ID (NO:2	7:					
Me	G GA t As	T GC	A AT	t Ly	G AG S Ar S	A GG(g Gl)	CT(L TGI	C TG' S Cy:	s Va	G CTI	G CT(s CT6	G TG L Cy:	T GGA S Gly	48
GC A l	A GT a Va	C TT L Ph	C GT e Va 2	l Se	G CC	C AGO	CAC Glr	GA/ Glu 21	u Ile	C CA1	GC(CGA Arg	7 TTC Pho 30	Arg	A AGA 9 Arg	96
GI	C GG y Gl	y Ar	A GT g Va 5	A GA	A AAI	G TTG S Lev	7 T T T T T T T T T T T T T T T T T T T	Val	C ACA	GTC Val	TAT	TA1 Tyr 45	Gly	GT/ Val	CCT Pro	144
GT: Va	G TG L Tr S	p Ly	A GA	A GC/ u Ala	A ACC	C ACC Thr 55	Thr	CTA Leu	Pho	TGT Cys	GCA ALE	Ser	GA1	GC1	AAA Lys	192
GC/ AL	a Ty	T GA	T AC	A GAD	GT/ Val	A CAT L His	AAT Asn	GTT Val	TGG Trp	GCC Ala 75	Thr	CAT His	GCC	TGT Cys	GTA Val 80	240
Pro	C AC	A GA	C CCC	AAC Asr 28) Pro	GLn	GAA Glu	GTA Val	GTA Val 90	Leu	GAA Glu	AAT Asn	GTA Val	ACA Thr 95	Glu	288
CA1 His	TT1	AA(ATO Met 100	: Trp	Lys	AAT Asn	AAC	ATG Met 105	Val	GAA Glu	CAG Gln	ATG Met	GLN 110	Glu	GAT Asp	336
Ile	: Ile	115	Leu	ı Trp	Asp	CAA Gln	Ser 120	Leu	Lys	Pro	Cys	Val 125	Lys	Leu	Thr	384
CCA Pro	130	Cys	GTT Val	ACT Thr	TTA	AAT Asn 135	TGC Cys	AAG Lys	GAT ASP	GTG Val	AAT Asn 140	GCT Ala	ACT Thr	AAT Asn	ACC Thr	432
ACT Thr 145	ASN	GAT Asp	AGC Ser	GAG Glu	GGA Gly 150	ACG Thr	ATG Met	GAG Glu	AGA Arg	GGA Gly 155	GAA Glu	ATA Ile	AAA Lys	AAC Asn	TGC Cys 160	480
TCT Ser	TTC Phe	AAT	ATC	ACC Thr 165	ACA Thr	AGC Ser	ATA Ile	AGA Arg	GAT Asp 170	GAG Glu	GTG Val	CAG Gln	AAA Lys	GAA Glu 175	TAT Tyr	528
GCT Ala	CTT	TTT	TAT Tyr 180	LYS	CTT	GAT Asp	GTA Val	GTA Val 185	CCA Pro	ATA Ile	GAT Asp	AAT Asn	AAT Asn 190	AAT Asn	ACC Thr	576
AGC Ser	TAT Tyr	AGG Arg 195	TTG	ATA Ile	AGT Ser	TGT Cys	GAC Asp 200	ACC Thr	TCA Ser	GTC Val	ATT	ACA Thr 205	CAG Gln	GCC Ala	TGT Cys	624
						CCA Pro 215				His						672
GGT Gly 225	TTT Phe	GCG Ala	ATT Ile	CTA Leu	AAG Lys 230	TGT . Cys	AAT Asn	GAT Asp	Ly3	ACG Thr 235	TTC Phe	AAT Asn	GGA Gly	Lys	GGA Gly 240	720

Pr	A TO	T AA	A AA 'S AS	7 GT(n Va 24!	l Se	C ACA	GT/	A CA	A TG n Cy: 250	5 Thi	A CA	T GG/ S Gly	A AT	7 AG	CCA Pro	768
ST Va	A GT	A TO	A AC r Th 26	r Gli	A CTO	CTG J Leu	CT/	AA AA ASI 265	n Gly	AG1	CT/	A GCA J Ala	GA/ G G L C 270	ı Glı	GAG Glu	816
GT Va	A GT l Va	A AT L I L 27	e Ar	A TC1 g Ser	GAC Asp	AAT Asn	710 Phe 280	: Thi	G AAG	AAT Asn	GC1 Ala	AAA B Lys 285	The	ATA Ile	ATA : Ile	864
ÇT Va	A CA L GL 29	n Le	G AA	A GAA S GIL	TCT Ser	GTA Val 295	GAA	AT1	AAT ASF	TGT Cys	ACA Thr 300	. Arg	CCC Pro	AAC Asr	AAC Asn	912
AA Asi 30	ח זה	A AG r Ar	A AA/ g Ly:	A AGT S Ser	ATA Ile 310	CAT	ATA Ile	GGA	CCA Pro	GGG Gly 315	Arg	GCA Ala	TTT Phe	TAT	ACT Thr 320	960
AC/ Th:	A GG/	A GA	A ATA	I ATA I I I e 325	Gly	GAT Asp	ATA Ile	AGA	CAA Gln 330	Ala	CAT	TGT Cys	AAC Asn	ATT Ile 335	AGT Ser	1008
AG/ Arg	A GC/ g Ala	A AA	A TGG S Trp 340) Asn	GAC Asp	ACT Thr	TTA Leu	AAA Lys 345	Gin	ATA	GTT Val	ATA Ile	AAA Lys 350	Leu	AGA Arg	1056
GAA	CAA J Glr	1 TT1 1 Pho 355	e Glu	AAT Asn	AAA Lys	ACA Thr	ATA Ile 360	Val	Phe	AAT Asn	CAC	TCC Ser 365	TCA Ser	GGA Gly	GGG Gly	1104
GAC Asp	Pro 370	, 611	ATT Ile	GTA Val	ATG Met	CAC His 375	AGT Ser	TTT Phe	AAT Asn	TGT Cys	GGA Gly 380	GGA Gly	GAA Glu	TTT Phe	TTC Phe	1152
TAC Tyr 385	Cys	AAT Asr	TCA Ser	ACA Thr	CAA Gln 390	CTG Leu	TTT Phe	AAT Asn	AGT Ser	ACT Thr 395	TGG Trp	AAT Asn	AAT ASA	AAT Asn	ACT Thr 400	1200
GAA Glu	GGG	TCA Ser	AAT Asn	AAC Asn 405	ACT Thr	GAA Glu	GGA Gly	AAT Asn	ACT Thr 410	ATC Ile	ACA Thr	CTC L e u	CCA Pro	TGC Cys 415	AGA Arg	1248
ATA Ile	AAA Lys	CAA Gin	ATT Ile 420	ATA Ile	AAC Asn	ATG Met	GTG Val	CAG Gln 425	GAA Glu	GTA Val	GGA Gly	Lys	GCA Ala 430	ATG Met	TAT Tyr	1296
GCC Ala	CCT Pro	CCC Pro 435	Ile	AGA Arg	GGA Gly	CAA Gln	ATT I le 440	AGA Arg	TGT Cys	TCA Ser	Ser	AAT Asn 445	ATT	ACA Thr	GGG Gly	1344
CTG Leu	CTA Leu 450	TTA	ACA Thr	AGA Arg	Asp	GGT (Gly (455	GT Gly	ATT Ile	AAT Asn	Glu	AAT Asn 460	GLY	ACÇ Thr	GAG Glu	ATC Ile	1392
TTC Phe 465	AGA Arg	CCT Pro	GGA Gly	Gly	GGA Gly 470	CAT / Asp /	lTG let	AGG Arg	Asp	AAT Asn 475	TGG Trp	AGA Arg	AGT Ser	Glu	TTA Leu 480	. 1440
TAT Tyr	AAA Lys	TAT Tyr	Lys	GTA Val 485	GTA . Val	AAA /	lle i	Glu	CCA Pro 490	TTA Leu	GGA Gly	GTA Val	Ala	CCC Pro 495	ACC Thr	1488
						GTG C	iln /				T GA	GCGG	CCGC		•	1532

- (2) INFORMATION FOR SEQ ID NO:28:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 507 amino acids
 - (B) TYPE: amino acid (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 . 10 15

Ata Val Phe Val Ser Pro Ser Glu Glu Ile His Ata Arg Phe Arg Arg 20 2530

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Glu Asn Val Thr Glu 85 90 95

His Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met Glu Asp 100 105 110

Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Thr Leu Asn Cys Lys Asp Val Asn Ala Thr Asn Thr 130 135 140

Thr Asn Asp Ser Glu Gly Thr Met Glu Arg Gly Glu Ile Lys Asn Cys 145 150 155 160

Ser Phe Asn lie Thr Thr Ser lie Arg Asp Glu Val Gln Lys Glu Tyr 165 170 175

Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn Asn Asn Thr 180 185 190

Ser Tyr Arg Leu Ile Ser Cys Asp Thr Ser Val Ile Thr Gin Ala Cys 195 200 205

Pro Lys Ile Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala 210 215 220

Gly Phe Ala Ile Leu Lys Cys Ash Asp Lys Thr Phe Ash Gly Lys Gly 225 230 235 240

Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro 245 250 255

Val Val Ser Thr Gin Leu Leu Leu Ash Gly Ser Leu Ala Giu Giu Giu 260 265 270

Val Val Ile Arg Ser Asp Asn Phe Thr Asn Asn Ala Lys Thr Ile Ile 275 280 285

Val Gln Leu Lys Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn 290 295 300

Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr

The Sty Gtu Ite Ite Gty Asp Ite Arg Sth Ata His Cys Ash Ite Ser 325 330 335 Arg Ala Lys Tro Asn Asp Thr Leu Lys Gin Ile Val Ile Lys Leu Arg 340 345 350 Glu Gin Phe Glu Asn Lys Thr Ile Val Phe Asn His Ser Ser Gly Gly 355 360 365 Asp Pro Glu Ile Val Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe 370 375 380 Tyr Cys Asn Ser Thr Gin Leu Phe Asn Ser Thr Trp Asn Asn Asn Thr 385 390 400 Glu Gly Ser Asn Asn fhr Glu Gly Asn fhr Ile Thr Leu Pro Cys Arg $405 \hspace{1.5cm} 405 \hspace{1.5cm} 410 \hspace{1.5cm} 415$ Ile Lys Gln Ile Ile Asn Met Val Gin Glu Val Giy Lys Ala Met Tyr 420 425 430 Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly
435 440 445 Leu Leu Leu Thr Arg Asp Gly Gly Ile Asn Glu Asn Gly Thr Glu Ile 450 455 460 Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu 465 470 480 Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr 485 490 495 Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys 500 505

- (2) INFORMATION FOR SEQ ID NO:29:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 amino acids (B) TYPE: amino acid

 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: Linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gin Arg Glu Lys Arg 1 5 10 15

5

What is claimed is:

- 1. A recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain_(W->X) point mutation, wherein X is an amino acid residue other than tryptophan.
- 2. The recombinant nucleic acid molecule of claim 1, wherein X is a valine residue.
 - 3. The recombinant nucleic acid molecule of claim 1, wherein the nucleic acid molecule is a DNA molecule.
- 15 4. The recombinant nucleic acid molecule of claim 3, wherein the DNA molecule is a plasmid.
- The recombinant nucleic acid molecule of claim 4, wherein the plasmid comprises the sequence of the plasmid designated PPI4-tPA.
 - 6. The recombinant nucleic acid molecule of claim 1, wherein the C4 domain is an HIV-1 $_{\rm LAI}$ gp120 envelope glycoprotein C4 domain.

- 7. The recombinant nucleic acid molecule of claim 6, wherein the mutant HIV-1 gp120 envelope glycoprotein is a mutant HIV-1_{LAI} gp120 envelope glycoprotein.
- 30 8. The recombinant nucleic acid molecule of claim 1, wherein the C4 domain is an HIV- $1_{\rm IR-FL}$ gp120 envelope glycoprotein C4 domain.
 - 9. The recombinant nucleic acid molecule of claim 8,

WO 94/22477

wherein the mutant HIV-1 gp120 envelope glycoprotein is a mutant HIV- 1_{JR-FL} gp120 envelope glycoprotein.

- 10. The mutant HIV-1 gp120 envelope glycoprotein encoded by the recombinant nucleic acid molecule of claim 1.
 - 11. A vaccine which comprises a therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of claim 10, and an adjuvant.

10

5

12. A method of treating an HIV-1-infected subject, which comprises immunizing the HIV-1-infected subject with the vaccine of claim 11, thereby treating the HIV-1-infected subject.

15

- 13. A vaccine which comprises a prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of claim 10, and an adjuvant.
- 20 14. A method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the HIV-1-exposed subject with the vaccine of claim 13, thereby reducing the likelihood of the HIV-1-exposed subject's becoming infected with HIV-1.

25

- 15. A method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the non-HIV-1-exposed subject with the vaccine of claim 13, thereby reducing the likelihood of the non-HIV-1-exposed subject's becoming infected with HIV-1.
- 16. A method of obtaining partially purified antibodies which specifically bind to the CD4-binding domain of

- HIV-1 gp120 envelope glycoprotein, which method comprises (a) immunizing a non-HIV-1-exposed subject with the vaccine of claim 13, (b) recovering from the immunized subject serum comprising said antibodies, and (c) partially purifying said antibodies, thereby obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein.
- 10 17. The method of claim 16, wherein the subject is a human.
 - 18. The partially purified antibodies produced by the method of claim 16.
- 15 19. A pharmaceutical composition, which comprises a therapeutically effective amount of the partially purified antibodies of claim 18, and a pharmaceutically acceptable carrier.
- 20 20. A method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of claim 19 effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.
- 21. A method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of claim 19 effective to reduce the population of HIV-1 in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.
- 22. A composition which comprises a prophylactically effective amount of the partially purified antibodies of claim 18, and a pharmaceutically acceptable carrier.

WO 94/22477 PCT/US94/03282

90

- 23. A method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises administering to the HIV-1-exposed subject a dose of the composition of claim 22 effective to reduce the population of HIV-1 in the HIV-1-exposed subject, thereby reducing the likelihood of the subject's becoming infected with HIV-1.
- 10 24. The method of claim 23, wherein the subject is a medical practitioner.
 - 25. The method of claim 23, wherein the subject is a newborn infant.

15

20

- 26. A method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1 as a result of exposure thereto during an incident wherein there is an increased risk of exposure to HIV-1, which comprises administering to the subject immediately prior to the incident a dose of the composition of claim 22 effective to reduce the population of HIV-1 to which the subject is exposed during the incident, thereby reducing the likelihood of the subject's becoming infected with HIV-1.
- 27. The method of claim 26, wherein the subject is a medical practitioner.

FIGURE 1

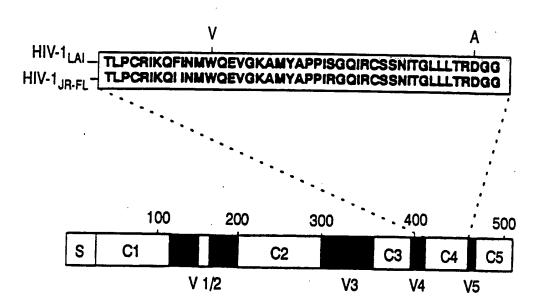


FIGURE 2

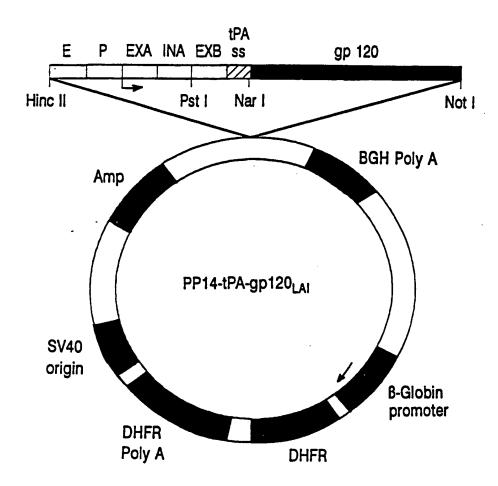


FIGURE 3D FIGURE 3E FIGURE 3F FIGURE 3F

-	ttgacattgattattgactagttattaatagtaatcaattacggggtcattagttcatagcccatatgga
73	gtteegegttaeataaettaeggtaaatggeeeggetgget
145	aataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacg
217	gtanactgeceacttggcagtacateaagtgtateatatgecaagtaegeeeeetattgaegteaatgaegg
289	taaatggcccgcctggcattatgcccagtacatgaccttatgggactttcctacttggcagtacatctacgt
361	attagtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggatagcggtttgactc
433	acggggatttccaagtctccaccccattgacgtcaatgggagtttgttt

FIGURE 3B

505

tecaaaatgtegtaacaaeteegeeeeattgaegeaatgggeggtaggegtgtaeggtgggaggtetatat aagcagagctcgtttagtgaaccGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAG **ACACCGGGACCGATCCAGCCTCCGCGGCCGGGAACGGTGCATTGGAACGCGGGATTCCCCGTGCCAAGAGTGA** ggccaacaccccgtcctagataggtgatggtatagcttagcctataggtgtgggttattgaccattattgac attggctatatgccaatactctgtccttcagagactgacacggactctgtatttttacaggatggggtccca ctccacgcgaatctcgggtacgtgttccggacatgggctcttctccggtagcggcggagctccacatccgag cctgtcccatgcccatgcctccagcggctcatggtcgctcggcagctccttgctcctaacagtggaggccag cactecectattggtgaegataettteeattaetaateeataaeatggeegetettgeeaeaaetatete tttattatttacaaattcacatatacaacgacgtccccgtgcccgcagtttttattaacatgcgggat acttaggcacaggacaatgcccaccaccaccagtgtgccgcacaaggccgtggcggtagggtatgtgtctga **★** Transcription Start Intron A Exon A 577 649 793 865 1009 721 937 1153 1225 1081

FIGURE 3C

aaatgagctcggagattgggctcgcaccgctgacgcagatggaagacttaaggcagcggcagaagaagatgc 1297

aggcagctgagttgttgtattctgtagagttggaggtaactcccgttgcggtgctgttaacggtggagggca 1369

gtgtagtctgagcagtactcgttgctgccgcgcgccaccagacataatagctgacagactaacagactgt 1441

teettteeatgggtetttetgeagTCACCGTCCTTGACACGATGGATGCAATGAAGAGAGGGCTCTGCTGT U tPA signal sequence ⋖ 1513

Nari > ட > H ,, 1585 11

agaacagaaaaattgtgggtcacagtctattatggggtacctgtgtggaaggaagcaaccaccactctattt ⋖ Н ш Ø Δ, 8 > ပ **A** 8 >4 > > H ⋖ > 3 O U **7** 1657

▲ Signal cleavage

tgtgcatcagatgctaaagcatatgatacagaggtacataatgtttgggccacacatgcctgtgtacccac > Z X, > Ы H Δ × 4 × 59 1729

gaccccaaccaagaagtagtattggtaatgtgacagaaattttaacatgtggaaaatgacatg Σ > z 83 1801

FIGURE 3D

Gaacagatgcatgaggatataatcagtttatgggatcaaagcctaaagccatgtgtaaaattaaccccact

1873

6/42 GTGCAGAAAGAATATGCATTTTTTTATAAACTTGATATAATACCAATAGATAATGATACTACCAGCTATACG **GGGGAAATGATGATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAG** TATTGTGCCCCGGCTGGTTTTGCGATTCTAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACA **TGTGTTAGTTTAAAGTGCACTGATTTGGGGAATGCTACTAATACCAATAGTAGTAATACCAATA**GTAGTAGC TTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACAT CTAGCAGAAGAAGAGGTAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAACCATAATAGTACAGCTG aatgtcagcacagtacaatgtacacatggaattaggccagtagtatcaactcaactgctgttgaatggcagt ທ ы z M ပ S ø Ŋ H H Ŋ 4 Δ, တ × z Z Z > H × > Δ H (Lu > H Z S Z S Q, ပ Z م O U H Ĺ 4 Δ 4 ပ æ z Z 0 a H z ¥ H × H G H J ပ ဟ × H H × တ M œ × H H Ω > U بعا H ഗ ပ > × ø > ပ Ш H G Δ Ш × > Σ z ω ပ ω Σ S ပ 1945 155 2089 179 2233 2305 2377 107 131 2017 2161 203 227 251

FIGURE 3E

8 0	A CA	MTC 1	X X	T T	₹ ×	R GA	NGA NGA
AACCAATCTGTAGAAATTAATTGTACAAGAACAACAATACAAGAAAAAGTATCCGTATCCAGAGGGGA N Q S V E I N C T R P N N T R K S I R I Q R G	CCAGGGAGACATTTGTTACAATAGGAAAATATGAGACAAGCACATTGTAACATTAGTAGAGCA P G R A F V T I G K I G N M R Q A H C N I S R A	AAATGGAATGCCACTTTAAAACAGATAGCTAGCAAATTAAGAGAACAATTGGAAATAATAAACAATAATC K w n a t l k q i a s k l r e q f g n n k t i i	CCTCAGGAGGGACCCAGAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTAC S S G G D P E I V T H S F N C G G E F F Y	CACAACTGTTTAATAGTACTTGGTACTTGGAGTACTGAAGGGTCAAATAACACT T Q L F N S T W F N S T W S T E G S N N T	GAAGGAAGTGACACACACCCCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAA E G S D T I T L P C R I K Q F I N M W Q E V G K	GCAATGTATGCCCCTCCCATCAGCGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGA A M Y A P P I S G Q I R C S S N I T G L L L T R	GATGGTGGTAATAACAACAATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGA D G G N N N N G S E I F R P G G G D M R D N W R
79 O	IGT.	Z F	rtt1 F	E Z	STAC V	Ϋ́L Y	Z Z
H H	HTTA	₹ ×	E	ည်း	E	TAT	ACA
GTA	Z Z	Z Z	ပ္သံ ပ		9 6 7)))	200 R
H I	GTJ	Z Z	S S	E	ეე ≽	ပ္ပံ ဖ	TG#
GTA	ATT H	S o	GTG C	r T	TGI M	T.	ATA
₹ _×	C C C	TTG F	ATT	GTA	S z	TTA	GAG
% ₩ %	SAG O	A O	TTA	€GGA ¥	H H	ATA	GAG G
₹	GAC	AAC	GIT	CII	TTA F	S SA	GAG G
ATA	TGA M	GAG R	ACA H	GTA	AAT	CAT	CTG
ACA	ATA N	T T	ည္မ	ATA	AAC K	GTT	GAC R
Z Z	8 0	AAT K	Ž,	TTA	TA I	gat R	TCA
	IAG.	J J	ITG I	SGT.	A GA	TTA I	TCT
3 4 C	3	A E	\$	CTT	5 .	\$ a	AGA, E
X	¥2.	₩	CAG	STAC	AT C	SAC	SCG.
STAC	§ .	CAS		ATAC	Ď,	ဗ္ဗ	3GT
ידדו י		3	75 E	71.	Ö T	ICA I	ATG
3	CTAC	3		iga.	֝֞֝֟֝֟֝֟֝֟֟֝֟֝֟֟֝֟֟֝֟֟֝֟֟֝֟֟֟֝֟֟֝֟֟֝֟֟֓֓֓֟֟֝֓֓֓֓֟֝֟֓֓֓֟֟֓֓	5	ACA.
3	19			Ω		71.	202
AG.	ATT	5			2	222	YTY I
TGT V	2004	95.	A10	7	163	7366	T.
ATC	GAG R	§ Z	AGCAA K 0	AATTCJ N S	24 S	GTA	95.
S o	Q	ATG W	TTTAAGCAAT	TGTAATTCAAC	ე ∀	AA A	17G
Žz	ပ္ပ	¥×	11	19 19 19	G B	ပ္ပ 🗸	g o
2449 299	2521 323	2593 347	2665 371	2737 395	419	2881	2953

FIGURE 3F

agtgaattat**aaatataaag**tagtaaaaattgaaccattaggagtagcacccaccaaggcaaagagaaga GTGGTGCAGAGAAAAATGAGOGGCCGC 3025 3097

FIGURE 4

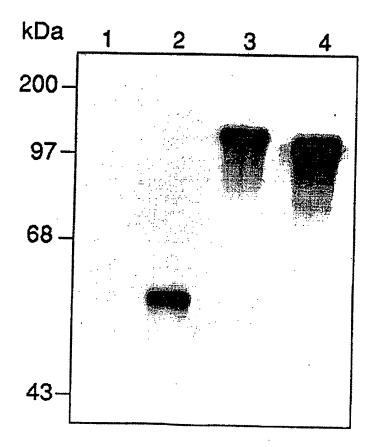


FIGURE 5A

Stable CHO	[gp120]
clone	(ng/ml)
5	6
6	14
9	123
10	4
12	18
13	18

FIGURE 5B

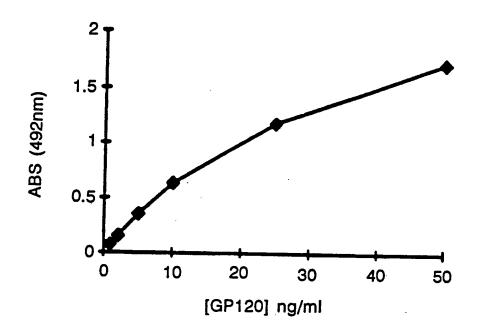


FIGURE 6

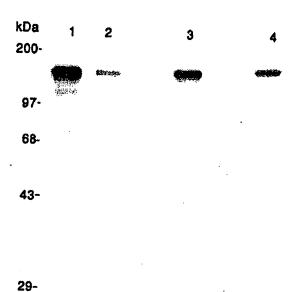


FIGURE 7A FIGURE 7A FIGURE 7B FIGURE 7B
--

FIGURE 7B

499 167	AGCATAAGAGATGAGGTGCAGAAAGAATATGCTCTTTTTTATAAACTTGATGTAGTACCA S I R D E V Q K E Y A L F Y K L D V V P	MG.	LGAT D	S E	616	350	Ž×	E E	TAT	ĞÇŢ ▼	CIT	TII	TAT	Z ×	ČŢ.	TGA D	TGT V	A G	S C	ర్ష
559 187	ATAGATAATAATACCAGCTATAGGTTGATAAGTTGTGACACCTCAGTCATTACACAG	TAA1 N	ZAT N	N N	D F	နိုင်ငံ	TAT	A 66	TTG	ATA	AGT S	1 61	9	TACK T	STC S	AGT.	CAT	TAC	2	9 4 0
619 207	GCCTGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT	ည်မ	N ×	ATA I	ည်င	TTT. F	SAG.	1 00 2 0	ATT	ပ္ပ	ATA	CA1	TAT	ည်ပ	ပ္ည	ن ا	ပ္ပ	95	744	T
679 227	GCGATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC A I L K C N D K T F N G K G P C K N V S	ICTA	×	ret C	X	GAT	X X	A CG	TIC	AAT	ဗွ် ဗ	Z ×	က် ပ	ည္က	ST. O	IX ×	\$z		5	ပ္ပမ္တ
739	ACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGCTGAATGGC T V Q C T H G I R P V V S T Q L L L N G	30	NTGT C	N F	CAT	ర్ట్ ల	ATT. I	A A	CCA P	GTA V	GTA V	STCA	S F	ည့်ဝ	KCT L	SCT.	. SCT	3	Q .	ပ္တပ္ဆ
799 267	AGTCTAGCAGAAGAAGGTAGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC S L A E E E V V I R S D N F T N N A K T	1 00 4 0	NGA.	E E	E E	GTA V	GTA.	ATT. I	AGA	TCT	GAC	AAI	TTC	3 CC	Ä	N N	ည္	5.	3	S F
859 287	ATAATAGTACAGCTGAAAATCTGTAGAATTAATTGTACAAGACCCAACAACAATACA I I V Q L K E S V E I N C T R P N N T	AGT.	CAG	ici 1	XX	GAA	ICT S	GTA V	E GA	ATT	AAT	រិក្ខិ	Z F	Z Z Z	ည္သ	A Z	g z	3_	YTA,	ទី ម
919	AGAAAAAGTATACATATAGACCAGGGAGAGCATTTTATACTACAGGAGAAATAATAGGA R K S I H I G P G R A F Y T T G E I I G	AAG1 S	IATA I	ICAT H	ATA	ပ္သိပ္သ	4 00	တ္တိ ပ	AGA R	GCA A	TTT F	TAT	. Y C	racu T	(S)	AGA. E	AAT	₹[PAG .	4 0
979	GATATAAGACAAGTTGTAACATTAGTAGAGCAAAATGGAATGACACTTTAAAACAG D I R Q A H C N I S R A K W N D T 1 K O	AAG. R	200	Çy ∢	CAT	ညီ ပ	Z Z	ATT	AGT S	AGA R	ည် 🔻	₹ ×	\TGC ¥	AX.	rgA(CAC T	TTT	3	5	. 9g

FIGURE 7C

								:
ATAGTTATAAAATTAAGAGAACAATTTGAGAATAAAACAATAGTCTTTAATCACTCCTCA I V I K L R E Q F E N K T I V F N H S S	GGAGGGGACCCAGAATTGTAATGCACAGTTTTAATTGTGGAGGAGAATTTTTCTACTGT G G D P E I V M H S F N C G G E F F Y C	AATTCAACACACTGTTTAATAGTACTTGGAATAATAATACTGAAGGGTCAAATAACACT NSTQLFNSTWNNTEGSNNTA	GAAGGAAATACTATCACACTCCCATGCAGAATAAAACAAATTATAAACATGTGGCAGGAA E G N T I T L P C R I K Q I I N M W Q E	GTAGGAAAAGCAATGTATGCCCTCCCATCAGAGGACAAATTAGATGTTCATCAAATATT V G K A M Y A P P I R G Q I R C S S N I	ACAGGGCTGCTATTAACAAGAGATGGTGTATTAATGAGAATGGGACCGAGATCTTCAGA T G L L L T R D G G I N E N G T E I F R	CCTGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATAAATATAAAGTAGTA P G G G D M R D N W R S E L Y K Y K V V	AAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGAAAAGAGAAA K I E P L G V A P T K A K R R V V Q R E	
ည်	¥ ×	Z Z	50	Z Z	TTC	GT3	A A	
CAC	TIC	Z Z	1GG ₩	ည္သ	ATC	≸ ×	80	
Z	FTT	ည် လ	Σ, Σ	ည် လ	3AG E	X	Si V	
F	E	ပ္ပ် ပ	IJz	ည် ည	ÿ ₽	\$ ×	F: >	
7 7 7 7	S S	EAG	Ž H	EA1	5 00	XXX	25 %	
TAG I	SAG G	CTG T	TTA I	ITA I	ATG N	IAI	₹	
\$	F.	T.	2	Ž	3	Y.	8	
5	ត្ត	¥ z	S _O	S O	GA	S a	Ž×	
\\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\	Z Z	Ž z	₹×	ပြွဲ ပ	\{\bar{\chi} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\	AG. 8	స్ట్ ∢	
A Z	TTI	A	ATA	A A	ATT	AGA R	₹	
SAG	AGT S	ენ 3	1 2 2 3	ZT I	S G T	ည် 🗷	N C	
17.17 17.17	S =	i H	ပ္ပံ့ပ	ប្លុស្ត	S S S	Z Z	ប្លុក្ខ	
Ž o	E Z	S	P G	ដ្ឋា	ATC	S S	5	
5	2	£_	ပ္ပို		9	g	9	
3	ទ្ធ	2 ×	2	ပ္မွ	5 4	S &	ج ال	
25 K	E H	F	Ž t	Y.	ប្តីម	ATC	ဦ် ပ	
ii.	3 a	CTO 1	ATC	ATG	TT.	GAT	TT.	8
\$ ×	ည်က	გი	FCT T	ည် န	CTA	ဗို့ ဗ	S a ;	AAATGAGGGGCGCC
I	DAG	Ž H	£ 2	Ž×	ii iii	S S S	AACCA E P	≧g
TT: V	ပ္တဲ့ ပွဲ့ ပ	ည္သ	ဥ ပ	2 0	ပ္ပဲ ဗ	1 50 0	TTG	GAG
8	A G	H.	8	8	3 6	1 6	\$	AT
AI	g o	\$ z	Б 🖽	E >	AC T	ပ္ပ 🏊	\$×	\$ ×
39	1099 367	59 87	219 407	279	339	399	459	519 507
1039	9	1159	1219	1279	1339	1399 467	1459	1519

.7

×

>

U

Δ,

×

Ы

S

O

Ω

E

H

<u> Agtttaaagtgcactgatttggggaatgctaataacaatagtagtaataccaatagtagtagtagcggggaa</u>

Z

z

G

H

H

S

133

397

16/42

FIGURE 8A

FIGURE 8A FIGURE 8C FIGURE 8B

LAI AV3

ATGGATGCAATGAAGAGAGGCCTCTGCTGTGTGCTG TCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCC aacccacaagaagtagtattggtaatgtgacagaaattttaacatgtggaaaatgacatggtagaacag **CTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAGGAAATCCATGCCCGATTCAGAAGAGGCGCCAGAACA** atgcatgaggatataatcagtttatgggatcaaagcctaaagccatgtgtaaaattaaccccactctgtgtt Signal cleavage 🛦 NarI U > 0 ပ G H Z ~ ĸ × H ¥ = 4 3 Σ H M Σ ⋖ ~ I × Z 0 Z Ξ ш > Z Δ, O W > H O > z × Ш > × > -1 တ H > > > ы Ω O Ы W z 109 37 13 37 181 253 85 325 109 61 SUBSTITUTE SHEET (RULE 26)

FIGURE 8B

469	ATGATG	ATG	GAG	3	9	GAG	ATA	X	Ş	ည်	TCT	TTC	Z	ATC	Ş	KCA.	S	ZZ.	IGA	GCT	AAG.	GTO	C.A.C.
157	M M M E K G E I K N C S F N I S T S I R G K V O	Σ	ω	×	U	M	H	×	z	ပ	တ	Ĺ.	z	н	S	₽	S	H	æ	U	×	} >	a
541	AAAGAA	TAT	ថ្ង	TT	TT	TAT	3	CTT	SATI	ATA	ATA	Z U	ATA(GAT	PAT(BATI	CI	Ş	ij	TAT	SC CG	TTG	AC A
181	KEYAFFYKLDIIPIDNDTTSYTLT	×	<	<u>fa</u>	ía,	>	×	1	۵	н	H	Δ,	H	۵	z	A	H	H	တ	×	H	ı	+
613	AGTTGT	*) NCC	ភ្ជ	GTC	ATT	2	S S	ည	ĮĞĮ	ర్ట	K AG	STA.	ည်	TTT(3 A GC	Š	VIIC	Ö	ATA	CAT	TAT	TGT
205	SCNTSVITOACPKVSFEPIPIHYC	Z	H	S	>	H	H	a	<	ပ	Δ,	×	>	S	(Eu	M	Ω,	н	Δ,	н	=	>	ပ
685	ອວວວວອ	ည	661	TT	Š	ATT	CIA	Ž	rgty	Z.	AAT.	SAS	SCG.	TIC	PAT(362	CAG	je AC	Š	IGI	ACA ACA	AAT	GTC
229	APAGFAILKCNNKTFNGTGPCTNV	4	U	ía,	<	H	H	×	ပ	z	z	×	H	Ĺ	z	ပ	H	U	Δ,	U	Ė	z	>
757	AGCACA	GIN	3	151	NC.	CAT	8 9	ATT	9	S S S	GTA	GTA:	₹	CI	CAAC	ST.	TG	TG	ZAT(360	AGT	CTA	Ş
253	STVOCTHGIRPVVSTQLLLNGSLA	>	0	ပ	H	=	ပ	H	~	م	>	>	တ	H	a	ı	ı	ı	z	ပ	S	ı	A
829	GAAGAA	GAG	GTA	GTA	LTA	AGA	ICI	7	₹ ET	TT CTT	ACA	GAC	¥ T€	3CT2	3	CC	VIA	TAG	3TA(C.A.G.	CTG	AA	CAA
277	E E E V V I R S A N F T D N A K T I I V Q L N Q	ш	>	>	H	æ	လ	4	z	Ĺ	H	۵	z	~	×	۲	H	н	>	o	1	z	a
901	TCTGTA	3	ATT	M	S	2	Ę	ij	7	1	TGT	Z A C	(TT)	ST ST	AGAC		7441	Ş	74	į V	į	Ę.	4
301	SVEINCTGAGECNISRAKWNATLK	M	H	Z	ပ	H	O	4	O	=	U	z	Н	တ	~	A	×	3	z	} ≪	į H	<u>د</u> د د	} ⊻

FIGURE 8C

973 325	CAGATAGCTAGCAAATTAAGAGAACAATTAGGAATAATAAACAATATTTAAGCAATCCTCAGGAGGG	AGC:	TAG	₹ ×	ATT	AAG R	AGA	ACA O	ATT' F	າີດີດີ. ດ	A Z	TAA.	Z K	AC.	caaattaagagaacaatttggaaataataaacaataatctttaagcaatcctcaggaggg : K L R E O F G N N K T I I F K Q S S G G	ATC I	TTT	'AAG K	ဦ ဝ	S S	STC.	1 257) (3 (3
1045 349	GACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACCACTGTTT D P E I V T H S F N C G G E F F Y C N S T Q L F	AGA	A AT	TGT. V	AAC T	S E	CAG S	ttt F	TAA	rrg	រ ិធិន្ស ខ	გ ეგ	3GAJ E	atti E	t tgtaacgcacagtittaattgtggagggaattitttttctactgtaattcaa cacagttt	TAC	TGT	AAT N	77. S	AC.	Z o	NCTC L	FTT:
1117	AATAGTACTTGGTTTAATAGTACTTGGAGGGTCAAATAACACTGAAGGAAG	TAC.	rtg ¥	STI	A Z	TAG	TAC.	777 8	3AG1 S	14 C3	IGA.	999	STC.	ZAA1	TTTAATAGTACTTGGAGTACTGAAGGTCAAATAACACTGAAGGAAG	ACT	GAA E	ဗ္ဗီ	AGT	GAC	ACA T	LATC H	ACA T
1189	CTCCCATGCAGAATAAAACATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCTCCCATC L P C R I K Q F I N M W Q E V G K A M Y A P P I	ÅT Ω	CAG. R	MAT! I	₹ ×	A o	ATT	IAT! I	Z Z	ATC	31G. *	20	E E	STA V	NTAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCTCCCATG I K Q F I N M W Q E V G K A M Y A P P I	A X	8 28	ATG	TAT	ပ္ပံ 🗸	ည် 🏲	ပ္ပ	ATC
1261	AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAACAAT S G Q I R C S S N I T G L L T R D G G N N N N	S O	E H	ragi R	ATG C	TTC S	ATC	A Z	[AT] I	T T	တ္တိ	CTC	E 1	ATT.	Gatgttcatcaaatattacaggctgctattaacaagagatggtggtaataacaacaa: R C S S N I T G L L L T R D G G N N N N	AGA	GAT	GGT G	G G 7	A Z	N N	Z Z	AAT N
1333	GGGTCCGAGATCTTCAGACCTGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAA G S E I F R P G G G D M R D N W R S E L Y K Y K	S E	AGATC E I	TTC	CAG.	ACC.	166 6	1 66	9	GAT	FATG	15	ο Ω	Z Z	TCAGACCTGGAGGAGATATGAGGGACAATTGGAGAGTGAATTATATAAATATAAA F R P G G G D M R D N W R S E L Y K Y K	AGA	AGT	GA.	TTA	TAT	¥ ×	TAT!	× Š
1405	GTAGTAAAAT V V K I	₹ ×	Ž H	S E	ပ္သို့ မ	ATT	ပြွ	AGT!	5 5 €	ည္က	.¥CC	.X.	ૄ	AAG *	TGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGAGAG	AGA M	GIG V	GIG V	9 c	AGA a	ES E	3 ×	TG.
1477	Not1	- 8 - 8											;	;	3	}	•	•	2	•	3	4 ,	

FIGURE 9A

JR-FL AV3

FIGURE 9B FIGURE 9C

FIGURE 9A

ATGGATGCAATGAAGAGA CATGCCCGATTCAGAAGAGGCGCAGAGTAGAAAGTTGTGGGTCACAGTCTATTATGGG gtacctgtgtggaaagaaaccaccactčtattttgtgčatcagatgctaaagcatat GATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCA Caagaagtagtattggaaaatgtaacagaacattttaacatgtggaaaaataacatggta gaacagatgcaggaggatataattaagtttatgggatcaaagcctaaagccatgtgaaa TTAACCCCACTCTGTGTTACTTTAAATTGCAAGGATGTGAATGCTACTAATACCACTAAT Σ ⋖ Ω Δ × Z S Signal cleavage Z 4 S ပ Z O Ω Ĺ, = 3 × > ഥ .. U ບ H S z 1 1 NarI 3 > H ~ > Z 띠 Z M 0 > × I W ပ U > a Δ 0 139 199 259 319 379 47 67 107 127 87

FIGURE 9B

439	GATAGCGAGGGAACGATGGAGAGAGAATAAAAACTGCTCTTTCAATATCACCACA D S E G T M E R G E I K N C S F N I T T
499 167	AGCATAAGAGATGAGGAGAAGAATATGCTCTTTTTTATAAACTTGATGTAGTACCA S I R D E V Q K E Y A L F Y K L D V V P
559 187	ATAGATAATAATACCAGCTATAGGTTGATAAGTTGTGACACCTCAGTCATTACACAGIO DI DIN NIN TISIN LISIC DI SIVII TO
619 207	GCCTGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT A C P K I S F E P I P I H Y C A P A G F
679 227	GCGATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC
739	ACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGCTAAATGGC T V Q C T H G I R P V V S T Q L L L N G
799 267	AGTCTAGCAGAAGAAGGTAGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC S L A E E E V V I R S D N F T N N A K T
859 287	ATAATAGTACAGCTGAAAATCTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAAC I I V Q L K E S V E I N C T G A G H C N
919 307	ATTAGTAGAGCAAATGGAATGACACTTTAAAACAGATAGTTATAAAATTAAGAGAACAA I S R A K W N D T L K Q I V I K L R E Q

FIGURE 9C

TTTGAGAATAAACAATAGTCTTTAATCACTCCTCAGGAGGGGACCCAGAAATTGTAATG F E N K T I V F N H S S G G D P E I V M	Cacagttttaattgtggaggaattttttttgtactgtaattcaacacactgtttaatagt h s f n c g g e f f x c n s t q l f n s	ACTTGGAATAATAATGCTGAAGTAACACTGAAGGAAATACTATCACACTCCCA T w n n n t e g s n n t e g n t i t l p	TGCAGAATAAAACATTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCT C R I K Q I I N M W Q E V G K A M Y A P	CCCATCAGAGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGAT PIRGOLIRCSSNITG	GGTGGTATTAATGAGAATGGGACCGAGATCTTCAGACCTGGAGGAGAGATATGAGGGAC G G I N E N G T E I F R P G G G D M R D	AATTGGAGAAGTGAATTATAAATAAAGTAGTAAAATTGAACCATTAGGAGTAGCA N W R S E L Y K Y K V V K I E P L G V A	Not
AACAATAGTCTTTAATCACTCCTC T I V F N H S S	TGTGGAGGAGAATTTTTCTACTG C G G E F F Y C	AATACTGAAGGGTCAAATAACAC N T E G S N N T	CAAATTATAAACATGTGGCAGGA	CAAATTAGATGTTCATCAAATAT Q I R C S S N I	GAGAATGGGACCGAGATCTTCAG E N G T E I F R	TGAATTATATAAATATAAAGTAG E L Y K Y K V	
TTTGAGAATAAA F E N K	CACAGITITAAI H S F N	ACTTGGAATAAT T W N N	TGCAGAATAAAA C R I K	CCCATCAGAGGA	GGTGGTATTAAT	AATTGGAGAAG' N W R S	
979 327	1039	1099 367	1159	1219	1279	1339	

FIGURE 10A

FIGURE 10A FIGURE 10B FIGURE 10C

LAI AV3-CD4"

ATGGATGCAATGAAGAGGGGCTCTGCTGTGTGCTG ctgctgtgtggagcagtcttcgtttcgcccagccaggaaatccatgccggattcagaagaggggcagaaca Nari ⋖

Signal cleavage ø ⋖ 囟 I Н ш Ø G ٩ S Þ ш > ۲ O O 109 37 13 37

tcagatgctaaagcatatgatacagaggtacataatgtttgggccacacatgcctgtgtacccacagaccc م > ပ 4 × H > Z × > M H Ω 4 181 61

aacccacaagaagtagtattggtaaatgtgacagaaaattttaacatgggaaaaatgacatggtagaacag Ω Z × 3 E 253

atgcatgaggatataatcagtttatgggatcaaagcctaaagccatgtgtaaaattaaccccactctgtgtt U O. H S O ۵ 3 .1 S H H Δ 325 109

agtttaaagtgcactgatttggggaatgctactaataccaatagtagtaataccaatagtagtagtgggggaa H Z Ø S Z H Z H z G H Ω 397 133

S

FIGURE 10B

469 157 541 181 ×

G

O

H

×

Z

Z

G

0

W

æ

×

ဟ

973 325

901 301

CAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATATCTTTAAGCAATCCTCAGGAGGG

23/42

ATGATGATGGAGAAAGGAGATAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAGGTGCAG aaagaatatgcatttttttataaacttgatataataccaatagataatgatactaccagctatacgttgaca **AGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACATTATTGT** GCCCCGGCTGGTTTTGCGATTCTAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACAAATGTC **Saagaaggtagtaattagatctgccaatttcacagacaatgctaaaaccataatagtacagctgaaccaa** a G > တ Н U H S H H Δ, H G Ω 回 × **~** တ z [a, Z K ۵ S S (L) z z H > H S Ω × ۵, × H ۵, z S ပ z ပ Ĺ 4 Z ۵ ر ح Z 4 × 0 .1 S × H U O **H** œ W H < > U S × ω

SUBSTITUTE SHEET (RULE 26)

685 229

829 277

757 253

613 205

FIGURE 10C

GACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTT x Ω 1045 349 1117

CTCCCATGCAGAATAAAACAATTTATAAACATGGTGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC Ö Ш H Z z S U ы H ഗ **Z** S Z 373 1189 397

Σ 4 × U > ы Ò > Σ H 0

AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC æ H H H O H Z S S ပ ø S 1261 421

gggtccgagatcttcagacctggaggaggagatatgagggacaattggagaaagtgaattatataaatataaa -1 М S **~** z ۵ æ Σ ۵ U G G Ω, æ 1333 445

٥, > G > 1405 469

NotI 1447 GOGGOGG

FIGURE 11A

FIGURE 11A FIGURE 11B FIGURE 11C

ATGGATGCAATGAAGAGA	M A M A M

JR-FL AV3-CD4"

۵ Ø > ш. > ⋖ G U H ы Þ U U 13

CATGCCCGATTCAGAAGAGGCGCAGAGTAGAAAAGTTGTGGGGTCACAGTCTATTATGGG H r r Nari 79 27

GTACCTGTGTGGAAAGAAGCAACCACCACTČTATTTTGTGČATCAGATGCTAAAGCATAT Δ တ Signal cleavage U (L) × 3 > 139 47

GATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCA Ω H ۵, > U 4 I H æ > Z > 199 67

Caagaagtagtattggaaaigtaacagaacattttaacatgtggaaaataacatggta Σ I Ш H > z ω O 259 87

gaacagatgcaggatataatcagtttatgggatcaaagcctaaagccatgtgtaaa > U ۵, ¥ S 0 Δ -1 S Ω 319 107

TTAACCCCACTCTGTGTTACTTTAAATTGCAAGGATGTGAATGCTACTAATACCACTAAT Z Z 4 Z > Δ × O Z H H > U 379 127

S O	GATAGCGAGGGAACGATGGAGAGAGGAGAAATAAAAAACTGCTCTTTCAATATCACCACA D S E G T M E R G E I K N C S F N I T T	GAG	4 00	₽ CG ∓	ATG.	GAG.	AGA R	ტე ე	GA.	ATA I	₹ ¥	Z	ည်ပ	rcr: s	ric. F	N	ATC	₽CC.	H F
ტ ა	AGCATAAGAGATGAGGAAGAATATGCTCTTTTTTAAACTTGATGTAGTACCA S I R D E V Q K E Y A L F Y K L D V V P	AGA R	GAT	GAG E	616 V	50	Ž×	GAA	TAT	GCT ▶	CIT	FTT	rati Y	×	LITT	SATO	STA(STA(P P
AI	ATAGATAATAATACCAGCTATAGGTTGATAAGTTGTGACACCTCAGTCATTACACAGID N N N T S Y R L I S C D T S V I T Q	AAT	N	MAT	1	ည်လ	rat. Y	A	TTG	ATA	AGT: S	ត្តិ	SAC	F F	CA(S) 	ATT)	ACA(S O
છું ≪	GCCTGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT A C P K I S F E P I P I H Y C A P A G F	ည္သိ 🏎	X	ATA	JCC: s	III(F	SAG E	ည်း	ATT	်ပ္သ	ATA(I	CAT	rati X	ត្តិ	300€	ည္က	3CI	GGT	ITT F
ပ္ပဲ 🕊	GCGATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC A I L K C N D K T F N G K G P C K N V S	CTA	X X	16	ZZ	SAT	\$ ×	ACG →	TTC	AAT	န ဗွ	Ž×	3GA(P	rg C	₹ ¥	Z Z	91C	နိုင္ငံင
N F	ACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGCTAAATGGC T V Q C T H G I R P V V S T Q L L L N G	30	TGT.	Ž to	CAT(ည် ပ	ATT	A	CCA	CTA(STAT	င် လ	ACTC	A O	73.50	74	T. T.	AAT	ပ္ပ ဗ
S	AGTCTAGCAGAAGAGAGGTAGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC S L A E E E V V I R S D N F T N N A K T	ర్ట్ ∢	SAN E	E GY	3AG(E	STA(STA V	ATT	AGA R	ICT(S	SAC	Z Z	ric.	CG.	Z Z	N	SCT.	₹	J L
AT	ATAATAGTACAGCTGAAAATCTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAAC I I V Q L K E S V E I N C T G A G B C N	GTA V	50	513	₹ ¥	SAA	က် လ	STA(SAN E	ATT	AAT	ဦ ပ	D F	55.0	9 <u>.</u> 4) () ()	E H	ည်ပ	Z Z
AT	ATTAGTAGAGCAAAATGGAATGACACTTTAAAACAGATAGTATAAAATTAAGAGAACAA ISRAKWNDILKQIVIKLREO	AGA R	ည္မွ	Ž×	1 662	AAT	3AC	ACT	TTA	X A	CAG	ATAC I	STTA	I	A	Ę,	AGA(3AAC E	₹ a

FIGURE 11C

TTTGAGAATAAACAATAGTCTTTAATCACTCCTCAGGAGGGGACCCAGAAATTGTAATG F E N K T I V F N H S S G G D P E I V M	CACAGITITAAITGIGGAGGAGAAITITICIACTGTAAITCAACACACIGIITAATAGI H S F N C G G E F F Y C N S I Q L F N S	ACTTGGAATAATACTGAAGGGTCAAATAACACTGAAGGAAATACTATCACACTCCCA T w n n n t e g s n n t e g n t i t d p	TGCAGAATAAAACAATATAAACATGGTGCAGGAAGTAGGAAAAGCAATGTATGCCCCT C R I K Q'I I N M V Q E V G K A M Y A P	CCCATCAGAGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGAT PIRG QIRC SSNITG LLLTRD	ggtggtattaatgagaatgggaccgagatcttcagacctggaggaggagatatgagggac g g i n e n g t e i f r p g g g d m r d	AATTGGAGAAGTGAATTATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCA N W R S E L Y K Y K V V K I E P L G V A	NotI CCCACCAAGGCAAAGAGAGAGTGGTGCAAAGAGAAAAATGAGCGGCCGC
ATAGICIT	GGAGGAGA G G E	ACTGAAGG T E G	ATTATAAA I I N	ATTAGATG I R C	aatgggac n g t	ATTATATA L Y	AGAAGAGT
AGAATAAAACA E N K T	STITIAATIGI S F N C	GAATAATAAT I N N N	AATAAAACAA	CAGAGGACAA R G Q	TATTAATGAG	GGAGAAGTGA W R S E	CAAGGCAAAG
TTTGA	CACAG H S	ACTIG T W	TGCAG C R	CCCAT	66166 6 6	AATT	CCCAC
979 327	1039	1099 367	1159 387	1219	1279	1339	1399

S

S

S

Z

H

Z

~

FIGURE 12A

FIGURE 12C FIGURE 12A FIGURE 12B

tPA signal sequence

LAI CD4

ATGGATGCAATGAAGAGAGGGCTCTGCTGT tgtgcatcagatgctaaagcatatgatacagaggtacataatgtttgggccacacatgcctgtgtacccaca gaccccaaacaagaagtagtattggtaatgtgacagaaaattttaacatggaaaaatgacatggta gaacagatgcatgaggatataatcagtttatgggatcaaagcctaaagccatgtgtaaaattaaccccactc tgtgttagtttaaagtgcactgatttggggaatgctactaataccaatagtagtaataccaatagtagtagc H Nari > U H Ω z L ĸ 3 ⋖ H Σ ပ 4 I × Z H بعا 3 × Ш > Z > -1 Q ۵, z ഥ S **100** I 0 H O > > Δ 3 8 W z ы > H G Ω L S H **>** > H 0 4 ⋖ > H H Signal cleavage × O Ω N C K O Ы H ы Σ ဟ 0 > 109 37 253 133 181 85 325 109 397 61

FIGURE 12B

29/42

GGGGAAATGATGATGGAGAAAGGAGAGATAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAG gtgcagaaagaatatgcattttttttataaacttgatataataccaatagataatgatactaccagctatacg ttgacaagttgtaacacctcagtcattacacaggcctgtccaaaggtatcctttgagccaattcccatacat tattgtgccccggctggttttgcgattctaaatgtaataataagacgttcaatggaacaggaccatgtaca **CTAGCAGAAGAAGAGGTAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAACCATAATAGTACAGCTG aaccaatctgtagaaattaattgtacaagacccaacaatacaagaaaaaa**gtatccgtatccagagggga CCAGGGAGAGCATTTGTTACAATAGGAAAATAGGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCA 1 1 S ρ, ы Ω G ပ H H a × တ Z (Le Z S I S 0 4 H × z H > S œ H O × Δ م ¥ > H œ بعا > z ဟ H ρ, Z H Σ Ų H م z z Z 4 æ Ω ပ G Z H a H ۵ × × H × ن æ H -1 × S × æ H H H W æ G > H ပ (Le, 4 H G H ပ Ĺų S (Le > Z H × ~ 0 H G H > > ~ M (d) H 띠 S 829 973 325 469 157 541 181 613 205 685 229 757 253 901 301

O

GTGGTGCAGAGAAAAATGAGCGGCCGC

1549

S

1477 493

469

· NotI

FIGURE 12C

AAATGGAATGCCACTTTAAAACAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAACAATAATC	TITAAGCAATCCICAGGAGGGGACCCAGAATIGIAACGCACAGITITAATIGIGGAGGGGAATITITCTAC	TGTAATTCAACACACTGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACT	gaaggaagtgacacacacacacacacacagaataaaacaatttataaacatggtgaggaagtaggaaaa	gcaatgtatgcccctcccatcagcggacaaattagatgttcatcaaatattacagggctgctattaacaaga	Gatggtggtaataacaatgggtccgagatcttcagacctggaggaggagatatgagggacaattggaga
Kwnaat kacataacaataatc	F K Q S S G G D P E I V I H S F N C G G E F F Y	C N S T Q L F N S T W F N S T W S T E G S N N T	e g s d t i t l p c r i k <u>q</u> f i n m v q e e v g k	A M Y A P P I S G Q I R C S S N I T G L L L T R	
AACAATTT E Q F	STTTTAAT S F N	CTTGGAGT	TTATAAAC I N	AAATATT	SAGGAGGA
TAAGAGI L R E	CGCACAC T H S	LATAGTAC N S 1	AACAATI K Q F	GTTCATC	GACCTGG
AGCAAA1	ATTGIAA	TGGTTTA	AGAATAA	ATTAGAI	ATCTTCA
S K	I V	W F	R I	I R	
ATAGCT! I A	CCAGAA	AGTACT: S T	CCATGC	GGACAAJ G Q	TCCGAG
AAAACAG	AGGGGAC	GTTTAAT	CACACTC	CATCAGO	CAATGGG
K O	G D	F N	T L	I S	
CACTIT	CTCAGG	CACAACT	CACAAT	CCCTCC	TAACAA
AATGGAATGCC	TTAAGCAATCC	AATTCAAC	AAGGAAGTGAC	CAATGTATGCC	GGTGGTAA
K w n a	F K Q S	N S T	E G S D	A M Y A	
1045 AAA	1117 TTT	1189 TGT	1261 GAA	1333 GCA	1405 GAT
349 K	373 F	397 C	421 E	445 A	

FIGURE 13A

FIGURE 13A FIGURE 13B

FIGURE 13C FIGURE 13D

JR-FL CD4"

ATGGATGCAATGAAGAGA

ဟ r Nari U 79 19

CATGCCCGATTCAGAAGAGGCGCAGAGTAGAAAAGTTGTGGGGTCACAGTCTATTATGGG Signal cleavage H

I

GTACCTGTGTGGAAAGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATAT 139 GATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCA Δ > > z X > W 199 67

CAAGAAGTAGTATTGGAAAATGTAACAGAACATTTTAACATGTGGAAAAATAACATGGTA Z Z × 3 Σ Z Ç., × > > 259 87

FIGURE 13B

\$ ×	TTAACCCCACTCTGTGTTAAATTGCAAGGATGTGAATGCTACTAATACCACTAAT L T P L C V T L N C K D V N A T N T T N	GATAGCGAGGGAACGATGGAGAGAGAATAAAAAACTGCTCTTTCAATATCACCACA D S E G T M E R G E I K N C S F N I T T	AGCATAAGAGATGAGGTGCAGAAAGAATATGCTCTTTTTTATAAACTTGATGTAGTACCA S I R D E V Q K E Y A L F Y K L D V V P	ATAGATAATAATACCAGCTATAGGTTGATAAGTTGTGACACCTCAGTCATTACACAGII DI NIN NITISIN NITIN NITISIN NITIN NITISIN NITIN N	TIT
GTA V	ACT	ACC.	GTA V	ACA T	G G
រក្សា	ACC F	AIC	GTA V	ATT	GCT A
ည် 🏎	Z	ZAAT	GAT	GTC V	ည္သ
Gaacagatgcaggaggatataatcagtttatgggatcaaagcctaaagccatgtgtaaaa E o m o e d i i s l w d o s l k p c v k	TACT	TTC	gcataagaggagggggaaagaatatgctcttttttataaacttgatgtagtacci s i r d e v q k e x a l f x k l d v v p	STC.	GCCTGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT A C P K I S F E P I P I H Y C A P A G F
CE T	<u>ر</u> ور م	STCJ S	\{\frac{1}{2}} \times	¥C	ည် က
AAG S	SAN N) (1)	rta) X	rgA(rta X
ည်တ	rgr(V	Ž z	ITT	ក្ស	Ž #
008 D	99 0	¥ ×	ICT.	AG: S	CAT
ATG W	25 ×	AAT	1 6€	GAT	700 P
TTT	110 0	AGA	ATA	GTT	MT
S S	XX	A 66	AGA	TAG	ပ္ပ
XI	111	GAG R	S. X	CTA	TGA
ITAT	TAC	455. 1	50	S S	CIT
199 S	31G1	CAT	1997 V	I I	CATO
9		S. J.	ATG.	ATY.	
D F		99	SAG.	ATA 1	37
A C	U U H	S CG	AT I	M T	10 C
AAC E	TAN.	ATA	S S	TAG	CCT A
u –	H	<u> </u>	4	4	O
319	379 127	439	499	559 187	619 207

FIGURE 13C

GCGATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC A I L K C N D K T F N G K G P C K N V S	CAGTA	ATCT	ATT	ু ব
T E	Ç m	15 ~	5	SAG *
G	ragge R	NATTAG I F	rgtagi V. E	AGGGAC G F
GATAA(D K	GGAAT	GTAGT	GAATC: E S	GGACC
GTAAT	CACAT T H	HAGAG E E	TGAAA L K	ATATA H
raaagt K	ATGTA	AGAAG	racago	STATAC
CGATTCI A I I	ACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGCTAAATGGC TVQCTLLLNG	AGTCTAGCAGAAGAGGGTAGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC S L A E E E V V I R S D N F T N N A K T	ATAATAGTACAGCTGAAAATCTGTAGAAATTAATTGTACAAGACCCAACAACAATACA I I V Q L K E S V E I N C T R P N N T	aglarabctatacatatagaccagggagagcattttatactacaggagaaataatagga R K S I H I G P G R A F Y T T G E I I G

FIGURE 13D

ATAGTTATAAAATTAAGAGAACAATTGAGAATAAAACAATAGTCTTTAATCACTCCTCA I V I K L R E Q F E N K T I V F N H S S	GGAGGGGACCCAGAATTGTAATGCACAGTTTTAATTGTGGAGAGAATTTTTCTACTGT G G D P E I V M H S F N C G G E F F Y C	AATTCAACACAACTGTTTAATAGTACTTGGAATAATAATACTGAAGGGTCAAATAACACT N S T Q L F N S T W N N T E G S N N T	gaaggaaatactaccacccatgcagaataaaacaaattataaacatggtgcaggaa e g n t i t l p c r i k q i i n m v q e	GTAGGAAAAGCAATGTATGCCCTCCCATCAGAGGACAAATTAGATGTTCATCAAATATT V G K A M Y A P P I R G Q I R C S S N I	ACAGGGCTGCTATTAACAAGAGATGGTATTAATGAGAATGGGACCGAGATCTTCAGA T G L L L T R D G G I N E N G T E I F R	CCTGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATAAATATAAAGTAGTA P G G G D M R D N W R S E L Y K Y K V V	AAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAGAG	
AATTAAGAGAAC. K L R E (CAGAAATIGTAA	AACTGTTTAATA(CIATCACACICCO I I I L	CAATGTATGCCC	FATTAACAAGAG L L T R 1	SAGATATGAGGG S D M R I	CATTAGGAGTAGG	1 2 00 0
ATAGTTATAA I V I	GGAGGGGACC G G D	AATTCAACAC N S T	GAAGGAAATA E G N	GTAGGAAAAG V G K	ACAGGGCTGC T G L	CCTGGAGGAG	AAAATTGAAC	NotI AAATGAGCGGCCGC K
1039	1099	1159 387	1219	1279	1339	1399	1459 487	1519

FIGURE 14A

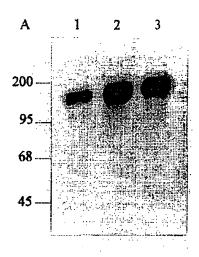


FIGURE 14B

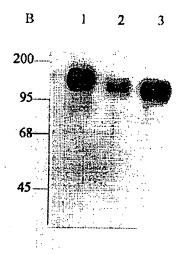


FIGURE 15A

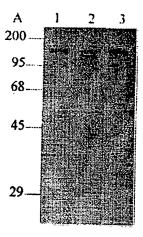


FIGURE 15B

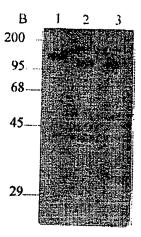
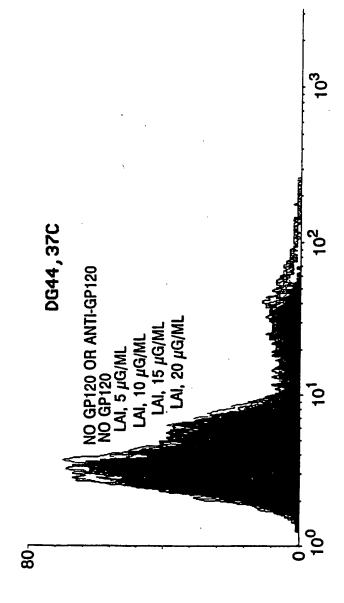


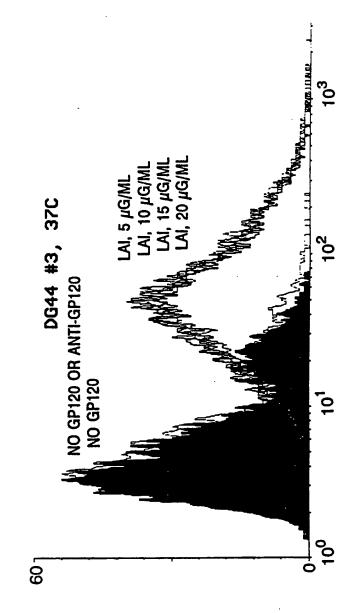
FIGURE 16A

39/42



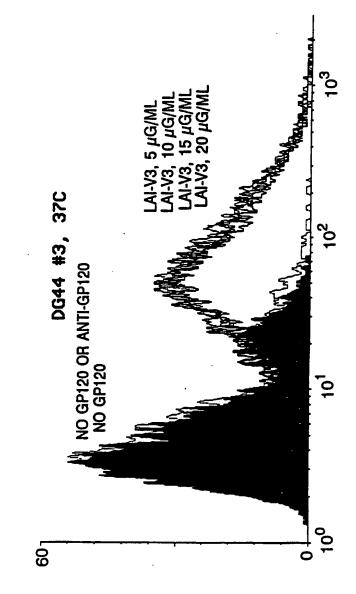
SUBSTITUTE SHEET (RULE 26)

FIGURE 16B

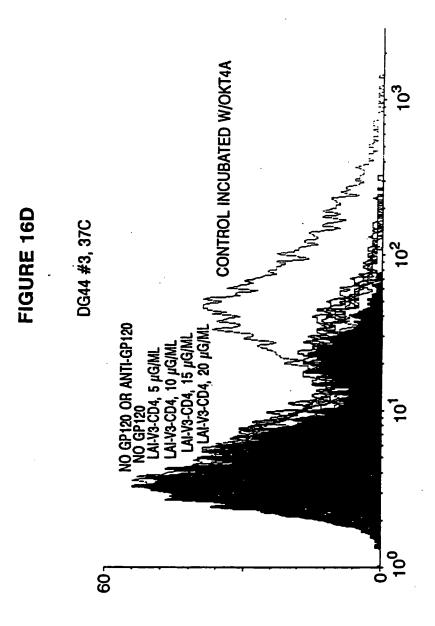


SUBSTITUTE SHEET (RULE 26)

FIGURE 16C



SUBSTITUTE SHEET (RULE 26)



SUBSTITUTE SHEET (RULE 26)

INTERNATIONAL SEARCH REPORT

International application No. PCT/US94/03282

IPC(5)	ASSIFICATION OF SUBJECT MATTER		
	:Please See Extra Sheet. : 424/88, 89; 536/27; 530/395		
According	to International Patent Classification (IPC) or to b	both national classification and IPC	
	ELDS SEARCHED		
	documentation searched (classification system folk	owed by classification symbols)	
	424/88, 89; 536/27; 530/395	ower by casemeanin symbols)	:
0.3.	424100, 89; 330/2/; 330/393		
Document	ation searched other than minimum documentation to	the extent that such documents an include	dim also Cald
		o and distances and successful documents are included	m the news searched
Electronic	data base consulted during the international search	(name of data base and subsequentian)	
	ialog, search terms: HIV-1, mutation, V3 loo	p, C4 region, envelope glycoprotein, V	accines, nucleic acid
C. DO	CUMENTS CONSIDERED TO BE RELEVANT	Γ	
Category*	Citation of document with individual		
oungo! y	Citation of document, with indication, where	appropriate, of the relevant passages	Relevant to claim No.
Y	Science, Volume 252, issued	17 May 1991 C Wein	
•	Hobson, et al, "LAV Revisited:	Origins of the Easts LIV 4	6,7
	Isolates from Institut Postour	Origins of the Early HIV-1	
	Isolates from Institut Pasteur*, article.	pages 301-305, see entire	
Y	IIS A F 020 440 (DEDZOFOK)		
•	US, A, 5,030,449 (BERZOFSKY)	ET AL) 09 July 1991, cols.	1-27
	47.	·	
v l	WO A 04/45540 105555		
']	WO, A, 91/15512 (GREGORY	ET AL) 17 October 1991,	1-27
j	entire patent.	ł	
,	1416	<u> </u>	
Y	WO, A, 91/11461 (PASEK ET	AL) 08 August 1991, see	1-27
	entire patent.		
		·	
ļ			
ł			
- 1			
		·	
Furthe	r documents are listed in the continuation of Box	C. See patent family annex.	
	ial estagories of cited documents:		
docu	ment defining the general state of the art which is not considered	date and not in conflict with the applicati	OR DUE CIZED TO sundersmoot the
	or particular relevance	principle or theory underlying the inven	tion
	er document published on or after the international filing date	"X" document of particular relevance; the considered sovel or cannot be considered	claimed invention cannot be
	ment which may throw doubte on priority claim(s) or which is to establish the publication date of another citation or other	when the document is taken alone	
4	et (casos (as abecnies)	"Y" document of particular relevance; the considered to involve an inventive at	laimed invention cannot be
docu Elican	ment referring to an oral disclosure, use, exhibition or other	comeaned with one or more other such d	ocuments, such combination
docu	ment published prior to the interestional filing date but later than	being obvious to a person skilled in the s	1
es pr	normy date caused		
E OI THE TO	tual completion of the international search	Date of mailing of the international search	h report
9 JUNE 19	94	JUN 2 4 1334	
	W		
ne and mai ommissione:	iling address of the ISA/US of Patents and Trademarks	Authorized officer	ma la
ox PCT eshington, I		LYNETTE F. SMITH	yza fr
simile No.			'
		Telephone No. (703) 308-0196	i

Form PCT/ISA/210 (second sheet)(July 1992)+

INTERNATIONAL SEARCH REPORT

International application No. PCT/US94/03282

A. CLASSIFICATION OF SUBJECT MATTER: IPC (5):	
A61K 39/12, 39/00; CO7K 17/00, 3/00, 13/00, 15/00; CO7H 15/12, 13/00	
<u> </u>	
·	
; ;	
,	